MOLECULAR IDENTIFICATION OF BACTERIA OF GENUS *STREPTOCOCCUS* AND RELATED GENERA

The present invention pertains to the area of diagnosis. More precisely, the invention concerns a method for the molecular identification of bacteria of genus *Streptococcus* and related genera *Enterococcus*, *Gemella*, *Abiotrophia* and *Granulicatella* using detection and/or amplifying and sequencing techniques with probes or oligonucleotide primers applied to strains of these bacterial genera.

Bacteria of the *Streptococcus* genus and of four related genera: *Enterococcus*, *Gemella*, *Abiotrophia* and *Granulicatella*, are Gram-positive and catalase-negative spherical bacteria of which more than around forty species are presently known. Bacteria of the genus *Lactococcus*, previously classified among the streptococci as Group N *Streptococcus*, do not come within the scope of this invention on account of their rare occurrence in human pathology, and because they can be easily distinguished from streptococci through their growth at +10°C. Genus *Streptococcus* officially comprises 55 species. Genus *Gemella* comprises 6 species, genus *Abiotrophia* comprises 1 species, genus *Granulicatella* comprises 3 species, and genus *Enterococcus* comprises 24 species [www.springer-ny.com/bergeysoutline/main.htm]. These species are easily and frequently cultured from environmental samples, veterinary clinical specimens and human clinical specimens [Ruoff Kl. (1999) in Manual of Clinical Microbiology, pp. 283-296, ASM Press]. In man, different species of the *Streptococcus* genus are responsible for community infections which may be severe

due to the invasive nature of the streptococci under consideration or through the production of possibly serious toxins with clinical signs distant from the site of infection. For example, *Streptococcus pyogenes* (Group A Streptococcus) is responsible for throat infections and post-streptococcal syndromes including rheumatic fever during which damage to the heart valves through an inflammatory process is responsible for possibly fatal heart valve disease. Also, several species of genus Streptococcus, in particular Group A, Group C and Group C Streptococci are responsible for life-threatening invasive infections, myositis in particular, i.e. degenerative changes to skin, subcutaneous and muscle tissue as has been described for some years. Also, *Streptococcus pneumoniae* (pneumococcus) for example causes pneumonia, meningitis and septicaemia. Bacteria of the genera *Streptococcus*, *Enterococcus*, *Gemella*, *Abiotrophia* and *Granulicatella* can cause endocarditis i.e. infection of the heart valves in man, which come under life-threatening infectious diseases [Casalta JP *et al.*, Journal Clinical Microbiology, 2002, 40: 1845-1847]. Also, some species of the genera under consideration can cause nosocomial infections, for example group A *Streptococcus* bacteria are responsible for bacteraemia subsequent to digestive endoscopy investigation. In addition, bacteria of the genus *Enterococcus* can cause nosocomial urinary infections after prophylactic antibiotic therapy with cephalosporins against which they are naturally resistant. These bacterial species also raise the problem of their increasing resistance to antibiotics, the resistance of *Streptococcus pneumoniae* to penicillin G [Garav J. Lancet Infect. Dis. 2002, 2: 404-415] and the resistance of *Enterococcus spp.* to vancomycin [Gold H.S., Clin. Infect. Dis. 2001, 33: 210-219; Bonten M.J. *et al.* Lancet Infect. Dis. 2001, 1 : 314-325].

These different bacterial species raise the problem of their detection in human pathological specimens and of their identification when isolated from such samples. Conventional detection methods rely on the evidencing of Gram-positive cocciform bacteria on direct examination of the pathological specimen. It is known, however, that this microscopic detection of bacteria of the genus *Streptococcus* and related genera in clinical specimens has a sensitivity threshold of 10^4 CFU/ml. It is therefore fully possible that a pathological specimen in man or animal contains one of the species under consideration which is not detected by direct microscopic examination of this pathological specimen. In addition, even though their structure is of Gram-positive bacterial type, they may give a false Gram-negative result after Gram staining of the pathological sample and give rise to erroneous or inconclusive identification. This is particularly frequent in bacteria of genus *Gemella*. In man, this is especially the case in anatomopathological and bacteriological investigation of the heart valves when diagnosing endocarditis.

When a bacterium of one of the species of the genera under consideration is isolated in the laboratory, conventional phenotype identification methods are the most commonly used to identify bacteria of species belonging to genus *Streptococcus* and related genera, and several identification kits and automated analysers have been developed to assist phenotype identification of bacteria of genus *Streptococcus* and related genera. In this respect, the extent of identification in routine practice is variable. In particular, one of the tests used for identifying Streptococci and bacteria of related genera is the detection of a haemolytic reaction, i.e. the destruction by the bacterium of red blood cells contained in a blood agar. However, this haemolytic reaction can be inhibited by the presence of oxygen

or by the presence of a peroxide when Streptococci bacteria are cultured in the presence of a high carbon dioxide concentration. Moreover, it is recognized that there exists a certain extent of subjectivity in assessing haemolysis by colonies of Streptococci and hence inter-operator variability which is detrimental to the quality of identification of these bacteria. For alpha-haemolytic streptococci, a second test is the optochin sensitivity test which enables identification of *Streptococcus pneumoniae* which is sensitive to this compound. However, strains of *Streptococcus pneumoniae* resistant to optochin have been reported [Lund E. Acta Patho. Microbiol. Immunol. Scand. 1959, 47, 308-315]. A final phenotype test is serotyping, which may also give false positive results in particular for streptococci in serogroup D on account of cross antigenicity between group D streptococci, *Enterococcus* and *Pediococcus*.

Several molecular systems have been developed to identify some serogroups or some species of genus *Streptococcus*, in particular for group A streptococci (*Streptococcus pyogenes*, *Streptococcus aginosus*, *Streptococcus constellatus*, *Streptococcus intermedius*) and group B (*Streptococcus agalactiae)* [Daly J.A. *et al*. J. Clin. Microbiol. 1991, 29:80-82; Heelan J.S. *et al*., Diagn. Microbiol. Infect. Dis. 1996, 24: 65-69] and for *Streptococcus pneumoniae* [Denys G.A. & Carrey R.B., J. Clin. Microbiol. 1992, 30: 2725 - 2727] by hybridisation of specific probes targeting the gene encoding the 16S ribosomal RNA. Also, different systems based on PCR amplification of genes coding for toxins or virulence factors have been developed to discriminate *Streptococcus pneumoniae* from among α-haemolytic Streptococci [Salo P. *et al.*, J. Infect. Dis. 1995, 171: 479-482; Morrisson K. *et al*. J. Clin. Microbiol. 2000, 38, 434-437; Kaijalainen T. *et al*. J. Microbiol. Meth. 2002, 51: 111-118], and for the detection of

Streptococcus agalactiae [Mawn J.A. *et al.* J. Clin. Pathol. 1993, 46: 633-636]. These different systems, however, only allow the identification of one or of a few species of genus *Streptococcus*.

An identification system for three species of streptococcus has been developed, based on amplification of the 16S-23S spacer [Forstman P. *et al.* Microbiology, 1997, 143, 3491-3500] but in this work identification was limited to only a few species of animal interest: *Streptococcus agalactiae*, *Streptococcus dysgalactiae* and *Streptococcus uberis*. Also, at the present time it is essential for laboratories to have 2 separate molecular targets for the detection and identification of streptococci to overcome the risks of molecular contamination inherent in the use of a single target.

Finally, no detection and identification system for *Streptococcus*-related genera has been developed, and more particularly for bacteria of the genera *Enterococcus*, *Gemella*, *Abiotrophia* and *Granulicatella*.

The inventors have shown in the present invention that the *rpoB* gene forms a genetic marker which can be used for the detection and specific identification of the bacterium of each species in genus *Streptococcus* and in 4 related genera: *Enterococcus*, *Gemella*, *Abiotrophia* and *Granulicatella*.

Although this gene has previously been shown to have use as a tool in bacterial identification of different bacterial genera, no publication mentions its use for identifying bacteria of genus *Streptococcus* and the four related genera, and the advantage of this gene's sequence for the identification of the said bacteria has in no way been suggested. On the contrary, a few partial sequences of the *rpoB* gene in a few species, available in GenBank, showed slight heterogeneity placing in doubt the advantage of this

gene as an identification tool for these bacteria. Finally, the inventors have developed a tool for the simultaneous identification of four bacterial genera, requiring the development of degenerate primers which could not be deduced from any of the *rpoB* sequences determined for each species.

More particularly, the present invention concerns nucleic acid sequences specific to the genus or to each species of genus *Streptococcus* and related genera whose nucleotide sequence is derived from the *rpoB* gene of the said bacteria.

According to Lazcano *et al.* [J. Mol. Evol. (1988) 27: 365-376] the polymerase RNAs are divided into two groups as per their origin, one consisting of the RNA- or DNA-dependent viral polymerase RNAs and the other consisting of the DNA-dependent polymerase RNAs of eukaryote or prokaryote origin (archaebacteria and eubacteria). The eubacterial DNA-dependent polymerase RNAs are characterized by a simple, conserved multimeric constitution denoted "core enzyme" represented by $\alpha\beta\beta'$, or "holoenzyme" represented by $\alpha\beta\beta'\sigma$ [Yura and Ishihama, Ann. Rev. Genet. (1979) 13: 59-57].

Numerous studies have evidenced the functional role, within the multimeric enzymatic complex, of the β subunit of the eubacterial polymerase RNA. Archaebacterial and eukaryote polymerase RNAs have a more complex structure possibly reaching ten and even thirty subunits [Pühlet *et al.* Proc. Natl. Acad. Sci. USA (1989) 86: 4569-4573].

The genes encoding the different $\alpha\beta\beta'\sigma$ subunits of the DNA-dependent polymerase RNA in eubacteria, the genes *rpoA*, *rpoB*, *rpoC* and *rpoD* respectively, are classified in different groups comprising the genes coding for constituent proteins of the ribosomal subunits or for enzymes involved in the replication and repair of the genome [Yura and Yshihma, Ann. Rev. Genet. (1979) 13: 59-97]. Some authors have shown that the sequences of the *rpoB* and *rpoC* genes could be used to

construct phylogenetic trees [Rowland *et al*. Biochem. Soc. Trans. (1992) 21 :40S] enabling separation of the different branches and sub-branches among the kingdoms of the living.

Before setting forth the invention in more detail, different terms used in the description and claims are defined below:

- By "nucleic acid extracted from bacteria" is meant either the total nucleic acid, or the genomic DNA, or the messenger RNAs, or the DNA obtained from reverse transcription of the messenger RNAs.
- A "nucleotide fragment" or an "oligonucleotide" are two synonymous terms designating a chain of nucleotide motifs characterized by an information sequence of the natural (or optionally modified) nucleic acids and able to hybridise, like natural nucleic acids, with a complementary or substantially complementary nucleotide fragment under predetermined conditions of high stringency. The chain may contain nucleotide motifs having a different structure to natural nucleic acids. A nucleotide fragment (or oligonucleotide) may for example contain up to 100 nucleotide motifs. It generally contains at least 8, and in particular at least 12 nucleotide motifs, further particularly 18 to 35, and may be obtained from a natural nucleic acid molecule and/or by genetic recombination and/or by chemical synthesis.
- A nucleotide motif is derived from a monomer which may be a natural nucleotide of a nucleic acid whose constituent elements are a sugar, a phosphate group and a nitrogenous base chosen from among adenine (A), guanine (G), uracil (U), cytosine (C), thymine (T); or else the monomer is a nucleotide modified in at least one of the three preceding constituent elements; as an example, modification may occur either at the bases, with modified bases such as inosine

which can hybridise with any base A,T,U,C or G, methyl-5-deoxycytidine, deoxyuridine, dimethylamino-5-deoxyuridine or any other modified base able to hybridise, or at the sugar, for example the replacement of at least one deoxyribose by a polyamide (Nielsen PE *et al.*, Science (1991) 254: 1497-1500], or at the phosphate group, for example through replacement by esters chosen from among diphosphates, alkylphosphonates and phosphorothioates.

- By "hybridisation" is meant the process during which, under suitable conditions, two nucleotide fragments having sufficiently complementary sequences are able to join together by stable, specific hydrogen bonds to form a double strand. Hybridisation conditions are determined by "stringency" i.e. the strictness of operating conditions. Hybridisation is more specific the higher the stringency. Stringency depends in particular upon the base composition of a probe/target duplex and on the extent of mismatch between two nucleic acids. Stringency may also be related to parameters of the hybridisation reaction, such as the concentration and type of ion species present in the hybridisation solution, the type and concentration of denaturing agents and/or the temperature of hybridisation. The stringency of the conditions in which a hybridisation reaction must be conducted depends in particular upon the probes used. All this data is well known and the suitable conditions may possibly be determined in each case by routine experiments. In general, depending upon the length of the probes used, the temperature for the hybridisation reaction lies between approximately 20 and 65°C, in particular between 35 and 65°C in a saline solution at a concentration of around 0.8 to 1 M.
- A "probe" is a nucleotide fragment having hybridisation specificity under determined conditions to form a

hybridisation complex with a nucleic acid having, in this case, a nucleotide sequence included either in a messenger RNA or in a DNA obtained by reverse transcription of said messenger RNA, the transcription product; a probe may be used for diagnosis purposes (capture and detection probes in particular) or for therapeutic purposes.

- A "capture probe" is a probe that is or may be immobilised on a solid support by any appropriate means, for example by covalency, adsorption, or direct synthesis on a solid. Examples of supports include microtitration wafers and DNA chips.
- A "detection probe" is a probe labelled with a marking agent chosen for example from among radioactive isotopes, enzymes in particular enzymes able to act on a chromogenous, fluorigenous or luminescent substrate (in particular a peroxidase or an alkaline phosphatase), chromophorous chemical compounds, chromogenous, fluorigenous or luminescent compounds, analogues of nucleotide bases and ligands such as biotin.
- A "species probe" is a probe enabling the specific identification of the species of a bacterium.
- A "genus probe" is a probe enabling the specific identification of the genus of a bacterium.
- A "primer" is a probe having 10 to 100 nucleotide motifs for example and having hybridisation specificity under determined conditions for enzymatic amplification reactions.
- By "amplification reaction" is meant an enzymatic polymerisation reaction, for example in an amplification technique such as PCR, initiated by primer oligonucleotides and using a polymerase DNA.
- By "sequencing reaction" is meant the obtaining of the sequence of a nucleic acid fragment or of a complete gene by means of an abortive polymerisation method using

oligonucleotide primers and said dideoxynucleotides [Sanger F, Coulson AR (1975), J. Mol. Biol. 94: 441] or multiple hybridisations with multiple probes fixed on a solid support such as used in DNA chips for example.

The sequences of the *rpoB* genes of the bacteria *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Streptococcus mutans* and *Streptococcus agalactiae* have been described in the literature.

The inventors have determined the complete sequences of the *rpoB* genes of other bacterial species of genus *Streptococcus* and related genera: *Streptococcus anginosus* and *Streptococcus equinus*, *Abiotrophia defectiva*, and a very large portion of the gene for *Streptococcus mutans* and *Enterococcus faecalis*. These species were chosen by the inventors as representing the main genetic groups determined on the basis of the study on the 16S gene in bacteria of genus *Streptococcus* and related genera, encompassing all the species currently described in this genus, so that the alignment of the *rpoB* sequences obtained in these species would most probably encompass all the *rpoB* sequences of all the species of these bacterial genera, more precisely they are therefore the most divergent species from a phylogenetic viewpoint among all the species currently described in this genus, so that the alignment of the *rpoB* sequences obtained in these species would most probably from a phylogenetic viewpoint encompass all the *rpoB* sequences of all the species of this bacterial genus.

From these complete or almost complete sequences, and after numerous unsuccessful attempts such as reported in examples 1 and 2 below, the inventors have evidenced the following consensus and specific sequences SEQ ID n°6 and 7:

- SEQ ID N°6: 5'- AARYTNGGMCCTGAAGAAAT-3', and
- SEQ ID N°7: 5'- TGNARTTTRTCATCAACCATGTG-3'

in which:

- N represents inosine or one of the 4 nucleotides A, T, C or G,
- R represents A or G,
- M represents A or C, and
- Y represents C or T,

and the reverse sequences and complementary sequences.

The inventors have determined said sequences SEQ ID n°6 and 7 as being not only consensual between all the bacteria of genus *Streptococcus* and said 4 related genera, but also specific to the family of bacteria of genus *Streptococcus* and said 4 related genera.

At the position corresponding to a nucleotide N,Y,M or R in sequences SEQ ID n°6 and 7, variable nucleotides are found in the complementary target sequences in relation to the species of the bacterium under consideration, but all the other nucleotides are conserved in all the species of bacteria of genus *Streptococcus* and of said 4 related genera.

Sequences SEQ ID n°6 and 7 so defined are present in the *rpoB* genes of all bacteria of genus *Streptococcus* and said 4 related genera, and are specific to the bacteria of genus *Streptococcus* and said 4 related genera, and can therefore provide genus probes or amplification primers to detect any bacterium of genus *Streptococcus* and of said 4 related genera.

For this purpose, one subject of the present invention is therefore an oligonucleotide which comprises a sequence of at least 8, preferably at least 12, further preferably between 18 and 35 nucleotide motifs, of which at least one sequence of 8, preferably 12, further preferably 18 consecutive motifs is included in one of the following sequences SEQ ID n°6 and 7:

- SEQ ID N°6: 5'-AARYTNGGMCCTGAAGAAAT-3', and
- SEQ ID N°7: 5'-TGNARTTTRTCATCAACCATGTG-3'

in which:

- N represents inosine or one of the 4 nucleotides A, T, C or G,
- R represents A or G,
- M represents A or C, and
- Y represents C or T

and the reverse sequences and complementary sequences.

Another subject of the invention is a mixture of oligonucleotides characterized in that it consists of an equimolar mixture of oligonucleotides of the invention, all having a different sequence and all comprising a sequence included in SEQ ID n°6 or all comprising a sequence included in SEQ ID n°7.

More particularly, a further subject of the invention is a mixture of said oligonucleotides, consisting of an equimolar mixture of 32 oligonucleotides of different sequences each comprising at least 15, preferably at least 18 consecutive nucleotide motifs included in the following sequence:

- SEQ ID n°6: 5' AARYTNGGMCCTGAAGAAAT-3'

in which:

- R represents A or G,
- Y represents C or T
- M represents A or C, and
- N represents A, T, C or G

and the reverse sequences and complementary sequences.

A further subject of the invention is a mixture of said oligonucleotides consisting of an equimolar mixture of 8 oligonucleotides having different sequences and each comprising at least 15, preferably at least 18 consecutive nucleotide motifs included in the following sequence:

- SEQ ID n°6: 5' AARYTNGGMCCTGAAGAAAT-3'

in which:

- R represents A or G,
- Y represents C or T

- M represents A or C, and
- N represents inosine

and the reverse sequences and complementary sequences.

A further subject of the invention is a mixture of said oligonucleotides, consisting of an equimolar mixture of 16 oligonucleotides having different sequences and each comprising at least 15, preferably at least 21 consecutive nucleotide motifs included in the following sequence:

- SEQ ID n° 7: 5′ TGNARTTTRTCATCAACCATGTG-3′

in which:

- R represents A or G, and
- N represents A, T, C or G

and the reverse sequences and complementary sequences.

A further subject of the present invention is a mixture of said oligonucleotides, consisting of an equimolar mixture of 4 oligonucleotides having different sequences and each comprising at least 15, preferably at least 21 consecutive nucleotide motifs included in the following sequence:

- SEQ ID n° 7: 5′-TGNARTTTRTCATCAACCATGTG-3′

in which:

- R represents A or G, and
- N represents inosine,

and the reverse sequences and complementary sequences.

Said mixtures of oligonucleotides are able to hybridise with a complementary sequence included in the *rpoB* gene of all the bacteria of genus *Streptococcus* and said 4 related genera, and can therefore be used as a genus probe or as amplification primers for the detection or respectively the amplification of an *rpoB* gene fragment of said bacterium.

To prepare said equimolar mixture of oligonucleotides using oligonucleotide synthesis methods known to persons skilled in the art, an equimolar mixture is used of 4 or 2

nucleotides for the nucleotides corresponding to N or respectively K,N,R or Y, namely:

- an equimolar mixture of the 4 nucleotides A, T, C and G for the nucleotides corresponding to N in which N represents A, T, C or G, and
- an equimolar mixture of the 2 nucleotides T and G for the nucleotides corresponding to K,
- an equimolar mixture of the 2 nucleotides A and C for the nucleotides corresponding to N,
- an equimolar mixture of the 2 nucleotides A and G for the nucleotides corresponding to R, and
- an equimolar mixture of the 2 nucleotides C and T for a nucleotide represented by Y.

Hence an equimolar mixture is obtained of 32 (2^3x4) and 16(2^2x4) nucleotides of different sequences for the 2 sequences SEQ ID n°6 and 7 respectively.

In said equimolar mixtures of oligonucleotides according to the invention, since "N" represents inosine which is able to hybridise with any base or an equimolar mixture of the 4 bases A, T, C, G, the sequences SEQ ID n° 6 and 7 are able to hybridise with the complementary sequence included in the *rpoB* gene of all bacteria of the *Streptococcus* genus and of the said 4 related genera.

In addition, these consensus sequences SEQ ID n°6 and n°7 flank hyper-variable sequences whose sequence is specific to each bacterium species of genus *Streptococcus*. These sequences flanked by SEQ ID n°6 and 7 may therefore be used as species probe for the bacteria of genus *Streptococcus* and said 4 related genera.

In addition, the sequences SEQ ID n°6 and 7 were determined as flanking an *rpoB* gene fragment comprising a zone whose variable length is approximately 720 bp and as

comprising the shortest sequences specific to each bacterium species of the *Streptococcus* genus and said 4 related genera.

The inventors were therefore able to evidence species probes for each of the examined 28 bacterial species of genus *Streptococcus* and said 4 related genera, corresponding to sequences SEQ ID n°8 to 35 described in example 2 below, flanked by the consensus sequences SEQ ID n°6 and 7.

A further subject of the present invention is a *rpob* gene or gene fragment of a bacterium of genus *Streptococcus* or of one of said 4 related genera, except sequences SEQ ID n°11, 12, 14, and of the bacteria *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans* and *Streptococcus agalactiae*, the reverse sequences and complementary sequences, characterized in that it comprises a sequence such as described in sequences SEQ ID n° 8 to 35 described in example 2.

A further subject of the invention is the complete sequence of the *rpoB* gene of the bacteria *Streptococcus anginosus*, *Streptococcus equinus* and *Abiotrophia defectiva* such as described in sequences SEQ ID n°1 to 3, which can be used in particular for a method of the invention.

A further subject of the present invention is the almost complete sequence of the *rpob* gene of the bacterium *Enterococcus faecalis* such as described in SEQ ID n°5, which can be used in particular for a method of the invention.

In sequences SEQ ID n° 1 to 3 and 5 and 8 to 35 described in the sequence listing at the end of the description:

- nucleotide M represents A or C,
- nucleotide K represents T or G,
- nucleotide R represents A or G,
- nucleotide W represents A or T,
- nucleotide Y represents C or T,
- nucleotide N represents A,T,C,G or I

- nucleotide S represents C or G,
- nucleotide V represents A,C or G

The consensus sequences derived from SEQ ID n° 6 and 7 evidenced according to the present invention, may be used as genus probe, as amplification or sequencing reaction primer in detection methods for bacteria of genus *Streptococcus* and said 4 related genera, by molecular identification.

With the sequences derived from sequences SEQ. ID n° 6 and 7 it is therefore not only possible to prepare genus probes for bacteria of genus *Streptococcus* and said 4 related genera, but also to detect and identify the species of said bacteria through amplification and sequencing using said sequences as primers.

The complete sequence of the *rpoB* gene may be used to identify the bacterium not only through the study of its primary sequence, but also through the study of the secondary and tertiary structures of the messenger RNA derived from transcription of the complete DNA sequence.

A further subject of the invention is an oligonucleotide or an *rpoB* gene fragment having a sequence included in or consisting of sequences SEQ ID n° 8 to 35, hence including sequences SEQ ID n° 11, 12, 14 and 22 of the bacteria *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans* and *Streptococcus agalactiae* respectively, and also among the oligonucleotides or fragments of reverse or complementary sequences such as defined above.

The inventors, after analysing the different sequences and comparing pair by pair all sequences SEQ. ID n° 8 to 35, determined that the homology rate between two different sequences among said sequences SEQ ID n° 8 to 35 is a maximum rate of 98.7% Below 98.7% homology between the sequences, they identify bacteria of different species. Consequently, a further subject of the invention is oligonucleotides or *rpoB*

gene fragments having sequences included in or consisting of said sequences SEQ ID n° 8 to 35, the reverse sequences, the complementary sequences and sequences showing at least 98.7% homology (i.e. a similarity rate of at least 98.7% between the sequences) with respect to said sequences SEQ ID n° 8 to 35, the reverse sequences and complementary sequences respectively.

The oligonucleotides, gene fragments and genes subject of the present invention have been described as comprising DNA sequences i.e. with nucleotides A, T, C and G. However, a further subject of the present invention is oligonucleotides comprising corresponding RNA sequences, i.e. in which T is replaced by U.

In the present description, by "reverse sequences and complementary sequences" is meant the following sequences:

- the reverse sequence of said sequence,
- the complementary sequence of said sequence, and
- the complementary sequence of the reverse sequence of said sequence.

Sequences SEQ ID n° 1 to 35 may be prepared by genetic engineering and/or chemical synthesis, in particular by automatic synthesis, using techniques well known to persons skilled in the art.

One first application of an oligonucleotide of the invention is its use a probe for the detection, in a biological specimen, of bacteria of one of the species of genus *Streptococcus* and said 4 related genera, which comprises a nucleotide sequence in one of the sequences SEQ ID n°6 to 35 and their reverse or complementary sequences.

An oligonucleotide comprising sequences SEQ ID n° 6 and 7 will be used as genus probe, and an oligonucleotide comprising a sequence included in or comprising one of sequences SEQ ID n° 8 to 35 will be used as species probe.

More particularly, the subject of the present invention is an oligonucleotide comprising a sequence specific to a bacterium species of genus *Streptococcus* and said related genera, preferably having at least 20 consecutive nucleotides, further preferably at least 30 consecutive nucleotides included in one of said sequences SEQ ID n° 8 to 35, or optionally an equimolar mixture of said oligonucleotides having different sequences.

Preferably, said sequences included in one of sequences SEQ ID n° 8 to 35, preferably having at least 20, further preferably at least 30 consecutive nucleotides included in one of said sequences SEQ ID n° 8 to 35, form the shortest sequences specific to the different respective species which can be used as species probe for *Streptococcus* bacteria and for said 4 related genera under consideration.

The probes of the invention may be used for diagnostic purposes, as mentioned previously, by determining the formation or non-formation of a hybridisation complex between the probe and a target nucleic acid in the specimen, using all known hybridisation techniques in particular "DOT-BLOT" techniques [Maniatis *et al.* (1982) Molecular Cloning, Cold Spring Harbor] DNA transfer techniques called "SOUTHERN BLOT" [Southern E.M., J. Mol. Biol. (1975) 98: 503], RNA transfer techniques called "NORTHERN BLOT", or so-called "sandwich" techniques in particular using a capture probe and/or a detection probe, said probes being able to hybridise with two different regions of the target nucleic acid, and at least one of said probes (generally the detection probe) being able to hybridise with a target region that is specific to the species, the capture probe and the detection probe evidently having nucleotide sequences that are at least partly different.

The nucleic acid to be detected (target) may be DNA or RNA (the first obtained after PCR amplification). When detecting a target of double strand nucleic acid type, the latter must first be denatured before starting detection. The target nucleic acid may be obtained using known methods for its extraction from a specimen to be examined. Denaturing of a double strand nucleic acid may be conducted using known chemical, physical or enzymatic methods, in particular by heating to an appropriate temperature, of over 80°C.

To implement the above-mentioned hybridisation techniques, in particular the "sandwich" techniques, a probe of the invention called a capture probe is immobilised on a solid support, and another probe of the invention called a detection probe is labelled with a marking agent. Examples of supports and marking agents are those previously given.

Advantageously, a species probe is immobilised on a solid support, and another species probe is labelled with a marking agent.

Another application of an oligonucleotide of the invention is its use as nucleotide primer comprising a monocatenary oligonucleotide chosen from among oligonucleotides having a sequence of at least 12 nucleotide motifs included in one of sequences SEQ ID n° 6 to 35, which can be used in the synthesis of a nucleic acid in the presence of a polymerase using a known method, in particular by amplification methods using said synthesis in the presence of a polymerase (PCR, RT-PCR, etc). In particular, a primer of the invention may be used for the specific reverse transcription of a messenger RNA sequence of a bacterial species of genus *Streptococcus* and said 4 related genera to obtain a corresponding complementary DNA sequence. Said reverse transcription may form the first stage of the RT-PCR technique, the following stage being PCR amplification of the

complementary DNA obtained. Primers of the invention may also be used for specific amplification, by chain polymerisation reaction, of the total DNA sequence of the *rpoB* gene of a species of genus *Streptococcus* and said 4 related genera.

In one particular case, said primer comprising an oligonucleotide of the invention also comprises the sense or antisense sequence of a promoter recognized by a polymerase RNA (promoters T7, T3, SP6 for example [Studier FW, BA Moffatt (1986) J. Mol. Biol. 189:113]: said primers can be used in nucleic acid amplification methods using a transcription step such as, for example, NASBA or 3SR techniques [Van Gemen B *et al*. Abstract MA 1091, 7th International Conference on AIDS (1991) Florence, Italy].

A further subject of the invention is a nucleotide primer comprising an oligonucleotide chosen from among oligonucleotides having a sequence comprising one of sequences SEQ ID n° 6 to 35 or a sequence included in SEQ ID n° 6 to 35 which can be used for full or partial sequencing of the *rpoB* gene of any strain of a bacterial species of genus *Streptococcus* and said 4 related genera.

Full or partial sequencing of the *rpoB* gene in any bacterium of genus *Streptococcus* and related genera enables the identification of all bacteria of genus *Streptococcus* and of said 4 related genera by bio-computerized analysis of this sequence, and enables the recognition of new unknown bacterial species of *Streptococcus* and of said 4 related bacteria.

Preferably, for use as a primer or for sequencing *rpoB* genes, said mixture of oligonucleotides of the invention is used, and further preferably said mixtures of oligonucleotides consisting of sequences SEQ ID n° 6 and SEQ ID n° 7.

More precisely, the present invention provides a detection method by identification to detect a bacterium of

one of the species of genus *Streptococcus* and of said 4 related genera, characterized in that the following are used:

- a complete or almost complete *rpoB* gene of said bacterium according to the present invention and an *rpoB* gene or gene fragment of a bacterium *Streptococcus pyogenes*, *Streptococcus pneumoniae*, *Streptococcus mutans* and *Streptococcus agalactiae* comprising a sequence such as described in sequences SEQ ID n° 11, 12, 14 and 22 respectively, the reverse sequences and complementary sequences, which may be used in particular as species probe, and/or
- a said fragment of said *rpoB* gene of said bacterium according to the present invention, comprising a nucleotide sequence chosen from among one of sequences SEQ ID n° 8 to 35, the reverse sequences and complementary sequences, which may be used in particular as species probe, and/or
- an oligonucleotide of the present invention comprising a sequence included in one of sequences SEQ ID n°8 to 35, the reverse sequences and complementary sequences, which may be used in particular as species probe, and/or
- an oligonucloetide or said mixture of oligonucleotides of the present invention comprising a sequence consisting of consecutive nucleotide motifs, included in one of sequences SEQ ID n° 6 and 7, which may be used in particular as genus probe or amplification primer.

Preferably, in said detection method of the invention, the following are used:

- a said *rpoB* gene fragment of said bacterium comprising a sequence chosen from among one of sequences SEQ ID n° 8 to 35 or an oligonucleotide having a sequence included in one of said sequences SEQ ID n° 8 to 35, the reverse sequences and complementary sequences, and/or

- at least one said mixture of oligonucleotides according to the present invention whose preferable sequences consist of sequences SEQ ID n° 6 and 7, and their reverse sequences and complementary sequences in which further preferably N represents inosine.

In a first embodiment of a detection method of the invention, it is sought to evidence the presence of a bacterium of genus *Streptococcus* and said 4 related genera, and the following steps are performed in which:

1. at least one genus probe comprising a said mixture of oligonucleotides having sequences comprising or included in one of sequences SEQ ID n° 6 and 7, the reverse sequences and complementary sequences according to the invention, is contacted with a specimen containing or possibly containing nucleic acids of at least one said bacterium of genus *Streptococcus* and of said 4 related genera, and
2. the formation or non-formation is determined of a hybridisation complex between said genus probe and the nucleic acids of the specimen, and the presence is determined of said bacterium of genus *Streptococcus* or of said 4 related genera if a hybridisation complex is formed.

In a second embodiment of a detection method for a bacterium of genus *Streptococcus* and said 4 related genera, the steps are performed in which:

1. Amplification primers, comprising said mixtures of oligonucleotides containing a sequence included in said sequences SEQ ID n° 6 and 7 reverse sequences and complementary sequences of the invention, are contacted with a sample containing or possibly containing nucleic acids of at least one said bacterium of genus *Streptococcus* and of said 4 related genera, using:

- as 5′ primer: a said mixture of oligonucleotides containing a sequence included in sequence SEQ ID n° 6 or preferably consisting of said complete sequence SEQ ID n°6, or a complementary sequence of the invention,
- as 3′ primer: a said mixture of oligonucleotides containing a sequence included in sequence SEQ ID n° 7 or preferably consisting of said complete sequence SEQ ID n°7, or respectively a complementary sequence of the invention.

2. The nucleic acids are amplified by enzymatic polymerisation reaction, and the occurrence or non-occurrence of an amplification product is determined, and in this way the presence is determined of said bacterium in the specimen if an amplification product is produced.

This second embodiment may be used to specifically detect the genus of a *Streptococcus* bacterium or said 4 related genera.

However, at step 2 of this second embodiment, it may be sought to specifically detect a given bacterium species of genus *Streptococcus* chosen from among the species *Streptococcus mutans*, *Streptococcus oralis*, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Streptococcus salivarius*, *Streptococcus sanguinis*, *Streptococcus suis*, *Streptococcus acidominimus*, *Streptococcus agalactiae*, *Streptococcus anginosus*, *Streptococcus constellatus*, *Streptococcus difficilis*, *Streptococcus dysgalactiae*, *Streptococcus equi*, *Streptococcus equinus*, *Streptococcus intermedius*, *Streptococcus mitis*, *Streptococcus bovis*, *Granulicatella adjacens*, *Abiotrophia defectiva*, *Enterococcus avium*, *Enterococcus casselliflavus*, *Enterococcus faecalis*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus sacharolyticus*, *Gemella haemolysins* and *Gemella morbillorum*,

as described in the variant of embodiment of a detection method specific to a species of said bacteria, given in the description below.

As previously set forth in the introduction, the genera *Streptococcus, Enterococcus, Granulicatella, Abiotrophia* and *Gemella* comprise more bacterial species than those effectively sequenced in this work. However, the sequenced species were chosen so that they encompass all known species in these bacterial genera and are sufficient in number to demonstrate the application of the *rpoB* sequence to the identification of the species of these genera.

In a variant of embodiment of a method of the invention for specifically detecting a species of said bacteria, the steps are performed in which:

1. a specimen containing or possibly containing nucleic acids of at least one said bacterium is contacted with at least one species probe consisting of said gene, said gene fragment or said oligonucleotide containing a sequence included in one of sequences SEQ ID n° 8 to 35, preferably an oligonucleotide consisting of one of said sequences SEQ ID n° 8 to 35, the reverse sequences and complementary sequences according to the invention, and
2. the formation or non-formation of a hybridisation complex is determined between said probe and the nucleic acids in the specimen, thereby determining the presence of said bacterium in the specimen if a hybridisation complex is formed.

In another variant of embodiment of the method of the invention, in which it is sought to specifically detect a given species of a bacterium of genus *Streptococcus* and of said 4 related genera, chosen from among the 28 species cited above, the method comprises the steps in which, in a specimen

containing or possibly containing nucleic acids of at least one said bacterium:

a) a sequencing reaction is conducted of an amplified *rpoB* gene fragment of said given bacterium using nucleotide primers consisting of said mixtures of oligonucleotides containing sequences included in sequence SEQ ID n° 6 as 5′ primer, and in SEQ ID n° 7 as 3′ primer, the sequences preferably consisting of said sequences SEQ ID n° 6 and 7, and their complementary sequences, and

b) the presence or absence of the given species of said bacterium is determined by comparing the obtained sequence of said fragment with the sequence of the complete *rpoB* gene of said bacterium or the sequence of a *rpoB* gene fragment of said bacterium containing said sequences n°8 to 35 and complementary sequences of the invention, thereby determining the presence of said bacterium in the specimen if the obtained fragment sequence is identical to the known sequence of the genus or of the *rpoB* gene fragment of said bacterium.

A further subject of the present invention is a diagnosis kit which can be used for a method of the invention, containing at least one said gene fragment or said oligonucleotide having a sequence included in or consisting of sequences SEQ ID n° 8 to 35, or a said oligonucleotide or mixture of oligonucleotides containing a sequence included in one of sequences SEQ ID n° 6 and 7, and/or at least one said *rpoB* gene fragment of said bacterium comprising sequences SEQ ID n° 8 to 35 and complementary sequences of the invention.

Advantageously, a kit of the present invention contains said oligonucleotides in the form of "biochips", i.e. fixed to solid supports, glass in particular, according to the method described in US patent 5,744,305 (Affymetrix, Fodor *et al*) using the resequencing strategy described in application WO

95/11995 (Affymax, Chee *et al*) or according to the method described by A. Troesch *et al.* in J. Clin. Microbiol., vol. 37(1), p 49-55, 1999. The oligonucleotides synthesized on the "biochip" carry out re-sequencing of the hyper variable region of the *rpoB* gene. This method offers considerable advantage in terms of production costs with no detriment to quality of identification of the different species through the choice of these identification sequences. Preferably, these oligonucleotides fixed onto the "biochip" solid support comprise 10 to 30 bases, e.g. 20 bases, with an interrogation position located in the central region for example at position 12 with respect to the 3' end of the sequence for oligonucleotides with 20 bases. Another example consists of using oligonucleotides having 17 bases with 2 interrogation positions: one at position 10 and one at position 8. Other oligonucleotides have lengths of between 10 and 25 nucleotides. The interrogation positions then vary according to the length of the oligonucleotide.

Analysis is conducted on the complete GeneChip® system (reference 900228, Affymetrix, Santa Clara, CA) which comprises the GeneArray® reader, the GeneChip® hybridisation oven, GeneChip® fluid station and GeneChip® analysis software.

An oligonucleotide of the invention may also be used as a gene therapy probe to treat infections caused by a strain belonging to a species of genus *Streptococcus* and said 4 related genera, said probe comprising an oligonucleotide such as defined previously. This gene therapy probe, able to hybridise on the messenger RNA and/or on the genomic DNA of said bacteria, may block translation and/or transcription and/or replication phenomena.

The principle of gene therapy methods is known and is based in particular on the use of a probe corresponding to an

antisense strand: the formation of a hybrid between the probe and the sense strand is able to disrupt at least one of the genetic information decoding steps. Gene therapy probes can therefore be used as anti-bacterial medicines, making it possible to fight against infections caused by bacteria belonging to the species of genus *Streptococcus* and said 4 related genera.

The invention will be more readily understood with the help of the description given below, divided into examples relating to experiments conducted with a view to implementing the invention and which are given solely for illustrative purposes.

Figure 1 shows the visualisation of the amplification products through ethidium bromide staining after electrophoresis on an agarose gel obtained in example 3.

Example 1: Sequence of the *rpoB* gene of three species of genus *Streptococcus* and related genera: *Abiotrophia defectiva*, *Streptococcus anginosus* and *Streptococcus equinus*.

The complete sequence of the *rpoB* gene of bacteria belonging to the species of *Abiotrophia defectiva*, *Streptococcus anginosus* and *Streptococcus equinus* was determined by enzymatic amplification and automatic sequencing available for Streptococci. The choice of these species was based on analysis of the 16S tree which shows genetic divergence covering the entire phylogenetic tree for streptococci.

Strategy and Sequencing:

Several partial 510-bp sequences of *rpoB* genes are available from GenBank for the 10 following streptococcus species: *Streptococcus intermedius*, *Streptococcus sanguinis*, *Streptococcus penumoniae*, *Streptococcus parasanguinis*, *Streptococcus oralis*, *Streptococcus mitis*, *Streptococcus*

cristalus, *Streptococcus constellatus*, *Streptococcus anginosus*, and *Granulicatell adjacens* [Majewski J., Zawadzki P., Pickerill P., Cohan F.M. and Dowson C.G. Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. J. Bacteriol. 182, 1016-1023 (2000)], but the primers used by these authors only amplify a fraction of the species of genus *Streptococcus*, and it was therefore not possible to carry out our work on the basis of this data alone. It was therefore necessary to determine primers able to amplify all strains of streptococci, enterococci, *Abiotrophia*, *Gemella* and *Granulicatella*. These primers also had to flank a region showing sufficient genetic diversity so as to be able to distinguish between two species. However, the alignment of these published partial sequences made it possible to determine the following common primers: (the numbering refers to the complete sequence of *Streptococcus pyogenes*)

SEQ ID n° 36: 5'- AGACGGACCTTCTATGGAAAA-3' (primer 748F)

SEQ ID n° 37: 5'- GGACACATACGACCATAGTG-3' (primer 116R), and

SEQ ID n° 38: 5'- GTTGTAACCTTCCCAWGTCAT -3' (primer 830R).

These primers allowed the sequencing of the central part of the *rpoB* gene with 714 bp for the five chosen species (*Streptococcus equinus*, *Streptococcus mutans*, *Streptococcus anginosus*, *Enterococcus faecalis*, and *Abiotrophia defectiva*. From this central fragment, sequencing was continued using the so-called genome Walker technique.

Outside this published zone [Majewski J. *et al*, J. Bacteriol. 2002, 182, 1016-1023], the alignment of the two complete sequences available from GenBank (*Streptococcus pneumoniae* [GenBank access number AE008542] and *Streptococcus pyogenes* [GenBank access number AE006480] made it possible to choose the following primers:

- SEQ ID n° 39: 5'- GTCTTCWTGGGYGATTTCCC-3' (primer 2215R)

- SEQ ID n° 40: 5'- ACCGTGGIGCWTGGTTRGAAT-3' (primer 2057R)
- SEQ ID n° 41: 5'- AACCAATTCCGYATYGGTYT-3' (primer 1252R)
- SEQ ID n° 42: 5'- AGIGGGTTTAACATGATGTC-3' (primer 371F)
- SEQ ID n° 43: 5'- AGIGCCCAAACCTCCATCTC-3' (primer 730F), and
- SEQ ID n° 44: 5'- CTCCAAGTGAACAGATGTGTA-3' (primer 585R)

With these primers, it was possible to extend the sequenced region for some of the five chosen strains. In fully unexpected manner, *E. Faecalis* is not amplified by these primers; but it was observed that the sequenced partial zone showed homology with the *rpoB* gene of *Listeria monocytogenes*, i.e. with a bacterium belonging to a different bacterial genus which could in no way be inferred from existing data, and we therefore chose primers in the *rpoB* gene of *Listeria* to amplify the *rpoB* gene of *Enterococcus faecalis*.

- SEQ ID n° 45: 5'-TTACCAAACTTAATTGAGATTCAAAC-3' (primer 180F)
- SEQ ID n° 46: 5'- AGTATTTATGGGTGATTTCCCA-3' (primer 410F)
- SEQ ID n° 47: 5'- GGACGTTATAAAATCAACAAAAAATT-3' (primer 910F)
- SEQ ID n° 48: 5'- AGTTATAACCATCCCAAGTCATG-3' (primer 2430R)
- SEQ ID n° 49: 5'- TGAAGTTTATCATCAACCATGTG-3' (primer 3280R)
- SEQ ID n° 50: 5'- CCCAAAACGTTGTCCACC-3' (primer 3360R)

The partial sequences so obtained for the five chosen strains (*Streptococus equinus*, *Streptococcus mutans*, *Streptcoccus anginosus*, *Enterococcus faecalis*, *Abiotrophia defectiva*) made it possible to choose the following primers:

- SEQ ID n°51: 5'- AACCAAGCYCGGTTAGGRAT-3' (primer 520R)
- SEQ ID n°52: 5'- ATGTTGAACCCACTIGGGGTGCCAT-3' (primer 2881F)

for the sequencing of the end C- and N- zones by Genome Walker.

Sequencing was then complete as displayed by the determination of the encoding region and the alignment of the translated proteins of the nucleotide sequences with the two published *rpoB* proteins of *Streptococcus pneumoniae* and *Streptococcus pyogenes*.

Several potential consensus primers were investigated to obtain a fragment able to lead to the complete sequence of the *rpoB genes* by successive elongations from a series of specific primers.

In each of the above steps, a large number of attempts with theoretically or potentially suitable primers failed before the above-mentioned primers were determined enabling the amplification and sequencing in successive steps of the entirety of the *rpoB* genes described below.

The sequencing reactions were conducted using reagents from the kit: ABI Prism dRhodamine Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer Applied Biosystems) in accordance with the manufacturer's recommendations and following the programme: 30 cycles comprising a denaturing step at 94°C for 10 sec., a hybridisation step of the primer at 50°C for 10 sec. and an extension step at 60°C for 2 minutes. The sequencing products were separated by electrophoresis on a polyacrylamide gel and 377 DNA sequencer (Perkin) and analysed to form consensus sequences using Sequence Assembler software (Applied Biosystems).

With this approach we were able to determine the complete sequence of the *rpoB* gene in two species of genus *Streptococcus* and in *Abiotrophia defectiva*:

SEQ ID n° 1: Sequence of the *rpoB* gene of *Streptococcus anginosus*. This sequence measures 4523 base pairs, has a guanosine plus cytosine content of 41% and is deposited in GenBank under accession number AF 535183:

5'-TCATACTTTTAGAGTCAGATTTAGCTGCTCTTTTTGTGCCTGTTTTGGGATTTTTGTCGTTTGT
CATCAAAATTAAAGATTCTGAAAATTACTCAAAAAGGATAAATGAAAATTGCTACTCTATTCCA
TTAATAGAGAATGTAGAAAGAAGAAGGAGTAAAAAACTTGGCAGGACATGAAGTTCAATACGGG
AAACACCGTACTCGTCGTAGTTTTTCAAGAATCAAGGAAGTTCTTGATTTACCAAATTTGATTG
AAATCCAGAGGATTCGTTCAAAGATTTTCTTGACCATGGTTTGAAAGAAGTATTTGAAGATGTA
CTTCCTATCTCAAACTTTACAGATACAATGGAGCTAGAGTTTGTTGGTTATGAAATTAAAGGAT
CTAAATACACTTTAGAAGAAGCACGTATCCATGATGCCAGCTATTCTGCACCTATTTTTGTGAC
TTTCCGTTTGATTAATAAAGAAACTGGTGAAATCAAAACCCAAGAAGTGTTCTTTGGCGATTTC
CCAATCATGACAGAAATGGGAACTTTCATTATCAATGGTGGTGAGCGGATTATCGTATCTCAGC
TCGTTCGTTCTCCAGGTGTTTACTTCAACGATAAAGTAGACAAAAATGGTAAAGTTGGTTATGG
TTCAACTGTCATTCCTAACCGTGGAGCTTGGTTAGAGCTGGAAACAGACTCAAAAGATATTGCT
TATACTCGGATTGACCGTACTCGTAAGATTCCGTTTACGACACTTGTTCGTGCGCTTGGTTTTT
CTGGCGATGATGAAATCTTTGACATTTTCGGCGACAGCGATCTCGTTCGCAACACGATTGAAAA
GGATATTCATAAAAATCCAATGGATTCACGTACGGATGAAGCGCTTAAAGAAATCTATGAACGT
CTTCGTCCAGGTGAGCCTAAAACAGCTGATAGTTCACGTAGTCTATTGGTCGCTCGTTTCTTTG
ATCCACATCGTTACGACTTGGCGGCAGTTGGTCGTTATAAAATCAATAAAAAATTAAACATTAA
AACACGTTTGTTAAATCAAACGATTGCAGAGCCTTTGGTAGATCCAGAAACAGGTGAAATCTTG
GTTGAAGCTGGAACGGTTATGACGCGTAGTGTCATTGATAGCATTGCAGAATACTTGGACGGTG
ATTTGAATAAATCACTTATATTCCAAATGATGCAGCTGTGTTAACAGAGCCAGTTGTTCTTCA
AAAATTCAAAGTGGTGGCGCCAACTGATCCAGATCGTGTGGTGACTATTATTGGTAATGCCAAC
CCAGGAGATCGAGTTCATACGATTACGCCAGCAGATATTTTGGCTGAGATGAATTACTTCTTGA
ACCTCGCTGAAGGACTTGGTCGTGTGGACGATATTGACCACTTGGGAAATCGTCGGATTCGTGC
CGTTGGTGAATTGCTTGCTAACCAAGTACGTCTTGGCTTGTCTCGTATGGAGCGAAACGTTCGG
GAGCGCATGAGTGTGCAAGATAATGAAGTGTTGACACCGCAACAAATCATTAACATCCGCCCAG
TCACAGCAGCTATCAAAGAATTCTTTGGTTCATCTCAATTGTCTCAATTTATGGACCAACATAA
TCCACTGTCTGAATTGTCTCACAAACGCCGTTTGTCAGCCTTGGGACCTGGTGGTTTGACTCGT
GATCGTGCTGGATATGAAGTGCGTGACGTGCACTATACCCACTATGGTCGTATGTGTCCGATTG
AAACGCCTGAAGGACCAAACATCGGTTTGATCAATAACTTGTCTTCTTATGGACACTTGAATAA
ATATGGCTTTATCCAAACGCCGTATCGTAAAGTGGATCGTGAAACAGGTCTGGTCACCAATGAA
ATCGTTTGGCTGACAGCGGACGAAGAAGATGAATTTATCGTAGCGCAAGCAAATTCTAAATTAA
CAGAAGATGGTCGTTTTGCAGAAGCGATTGTCATGGGACGTCACCAAGGGAACAACCAAGAATT
TCCTTCAGATCAAGTAGACTTCATGGATGTATCGCCTAAGCAGGTAGTTGCGGTTGCGACAGCA
TGTATTCCTTTCCTTGAAAACGACGACTCAAACCGTGCTCTCATGGGTGCCAACATGCAACGTC
AGGCGGTACCGTTGATTGATCCGCATGCACCATATGTTGGTACTGGTATGGAATACCAAGCAGC
TCATGACTCTGGTGCGGCGATTATTGCCCAACACGACGGTAAAGTTGTATATTCTGATGCAGCC
AAAGTTGAAGTTCGTCGTGAAGATGGCTCACTTGATGTCTATCATATTACGAAATTCCGCCGTT
CAAACTCTGGTACTTCTTACAACCAACGTACGCTGGTAAAAGTTGGCGATACAGTTGAAAAAGG
TGACTTTATCGCAGACGGACCTTCTATGGAAAAAGGTGAAATGGCACTTGGACAAAATCCAATC
GTTGCTTATATGACATGGGAAGGTTACAACTTTGAAGATGCCGTTATCATGAGTGAGCGTTTAG
TGAAAGACGATGTTTACACATCTGTTCACTTGGAGGAATTTGAATCAGAAACACGTGATACAAA STRF
GCTTGGACCTGAAGAAATCACGCGCGAAATTCCAAACGTCGGTGAAGATGCTTTGAGAGACCTT
GACGAAACGGGAATTATCCGCATTGGTGCTGAGGTAAAAGAAGGCGACATTCTTGTCGGTAAAG
TAACACCGAAAGGTGAAAAAGACTTATCTGCTGAAGAACGCCTGCTTCATGCAATTTTCGGTGA
TAAATCTCGTGAAGTACGTGATACTTCCCTTCGTGTACCACATGGTGGTGCAGGGGTTGTCCGT
GATGTGAAAATCTTTACTCGTGCGAACGGTGATGAATTGCAATCTGGTGTCAACATGTTGGTAC
GTGTTTACATCGCTCAAAAACGGAAAATCCGTGTTGGGGATAAGATGGCTGGACGTCACGGAAA
CAAAGGGGTTGTTTCCCGCATTGTTCCAGTTGAGGATATGCCGTATCTTCCAGATGGAACACCA

GTTGATATTATGTTGAACCCACTTGGGGTGCCATCTCGTATGAATATTGGTCAAGTTATGGAGC
TTCACCTCGGTATGGCTGCTCGCAACCTTGGCATTCACATTGCAACACCAGTATTTGACGGGGC
TAGCTCAGATGATCTTTGGGAAACCGTTCGTGAAGCTGGCATGGATAGCGATGCTAAGACAATC
CTTTATGATGGCCGTACTGGTGAGCCATTTGATAATCGTGTATCCGTTGGTGTCATGTACATGA
TCAAACTCCACCATATGGTTGATGATAAGCTCCATGCCCGTTCCGTTGGTCCTTATTCAACCGT STRR
TACGCAACAACCTCTTGGTGGTAAAGCGCAGTTTGGTGGACAACGTTTTGGAGAAATGGAAGTT
TGGGCTCTTGAAGCCTACGGTGCTTCTAACGTCCTTCAAGAAATCTTGACTTACAAGTCAGATG
ACATCAATGGTCGTTTGAGAGCTTATGAAGCCATTACCAAAGGTAAGCCAATTCCAAAACCAGG
TGTTCCAGAATCCTTCCGTGTCCTTGTAAAAGAATTGCAATCACTTGGTCTTGACATGCGTGTC
CTTGATGAAGACGACAATGAAGTCGAACTTCGTGACTTGGACGAAGGCATGGATGATGATGTGA
TTCATGTAGACGATCTTGAAAAAGCACGTGAAAAAGCAGCACAAGAAGCAAAAGCCGCTTTTGA
TGCTGAAGGGAAAGAATAAGAACTGATTCAATAGATAATAAAGAAAGGTAAGAAATAGTGGTTG
ATGTAAATCGTTTTCAAAGTATGCAAATCACCCTAGCTTCTCCTAGTAAAGTCCGCTCTTGGTC
TTATGGAGAAGTGAAGAAACCTGAAACAATTAACTACCGCACACTAAAACCAGAACGCGAAGGG
CTTTTTGATGAAGTCATCTTTGGTCCTACGAAAGACTGGGAATGTGCGTGTGGAAAATATAAAC
GGATTCGTTATAAAGGAATCATTTGTGACCGTTGTGGTGTTGAAGTAACTCGTACTAAAGTTCG
TCGTGAACGTATGGGACATATTGAGTTGAAAGCCCCAGTCTCCTCATATTTGGTATTTTAAAGG
AATTCCAANTCGCATGGGCTTGACCTTGGACATGAGCCCTCGTGCTCTTGAAGAAGTCATNTAN
TTTGCAGCTTATGTGGTGANTGACCCTAAAGATACNCCACTTGAGCACAAATCCATTATGACAG
AGCGGGATGGTTNGTGAACGCTGACNTGAATATGGCCAAGGCTCTTTTGTTGCAAAAATGGGTG
YTGAAGCAATCCAAGATCTNNTGAAACANGTAGACCTGGAAAAAGAAATTGCAGAGCTCAAAGA
TGAATTAAAAACGGCAAGTGGGCAAAAGCGCGTAAAMGCTAANTTCGTCGNTNNGACTCTTTTC
GATNCTTTCCAAAAATCATGGTACACAAAACCAGAACTGGATGGTCTTAAACCATCNTNTCACC
GCTCATTCCAGACAC -3'

SEQ ID n° 2: Sequence of the *rpoB* gene of *Streptococcus equinus*. This sequence measures 4118 base pairs, has a guanosine plus cytosine content of 41% and is deposited in GenBank under number GenBank accession AF 535187:

5'-CACGCGTGGTCGACGGCCCGGGCTGGTGAATTGTCATAAGTTGTGTAGTAGTAAATTCCCTTAT
CAGTGTTGATGCATGAGCTATAAATAGTGTACTCATATTTGCCACTTTCATCGACATAGCAAAG
TCCTTTTTGTTGTTCAACGGATTTTAAAATGTGGAAGAATTGATTAACACTGCTTTCTTCTGTT
TCTTCAGCCACAGAATTTAATTTTGTAAAAGTAACTTTTACATAACGTGACATTGATGATAAAT
CACCAGGCAAGCCAAGTCCACCCATGCCACGGCTATAAGTTTCAAGTTCTAACTCTTTAGCAAA
ACGATTTTCTGAAACCTTTGGAGATAGATGACGATAGTTATTCAAATTGAATAATTGTTTATCA
AAAGTTGGATTATTAGTCAAAACACCTGTTGAGTTATTCGTAAACTTATAGGGCACGCGTGGTC
GACGGCCCGGGCTGGTAAAGACTTCTTGGATAACGGATTAAMAGAAGTTTTTGAAGATGTACTT
CCGATTACAAACTTTACGGATACTATGGAGCTTGAATTTGTTGGTTACGAATTGAAAGAGCCTA
AGTATACGCTTGAAGAAGCTCGTATCCACGATGCATCTTATTCAGCACCTATTTTTGTAACCTT
CCGTTTGATTAATAAAGAAACAGGAGAAATCAAAACTCAAGAAGTTTTCTTCGGTGATTTCCCA
ATTATGACTGAAATGGGTACATTCATCATCAACGGTGGTGAACGTATTATCGTTTCTCAGTTGG
TTCGTTCTCCTGGTGTTTATTTCAACGATAAAGTTGATAAAAACGGTAAAGTTGGTTACGGTTC
AACTGTAATCCCTAACCGTGGAGCATGGCTTGAATTAGAAACAGATTCAAAAGATATTGCTTAC
ACACGTATCGACCGTACACGTAAAATTCCATTTACAACTCTTGTACGTGCGCTTGGTTTCTCAG
GTGATGATGAAATCATGGATATCTTTGGTGATAGCGAACTTGTTCGTAACACAATCGAAAAAGA
TATTCACAAAAACCCAGCAGACTCACGTACTGACGAAGCTCTTAAAGAAATTTACGAACGCCTT
CGTCCAGGTGAACCAAAAACAGCTGATAGCTCACGTAGCTTGCTTGTAGCTCGTTTCTTTGACC
CACGTCGTTATGACTTGGCAGCTGTTGGTCGTTACAAAATCAACAAAAAACTTAACATCAAGAC
TCGTCTTTTGAACCAAACAATCGCTGAAAACTTGGTTGATGCTGAAACTGGTGAAATCCTTGTT
GAAGCTGGTACAGTAATGACACGTGACGTGATTGATTCAATCGCTGATCAATTGGATGGTGACC
TTAACAAATTTGTTTACACACCAAATGATTACGCTGTTGTCACTGAACCTGTTGTTCTTCAAAA
ATTCAAAGTTGTTGCACCAAACGATCCAGACCGCGTTGTTACAATCGTTGGTAACGCAAATCCT
GATGACAAAGCGCGTGCGCTTACACCAGCTGATATCTTGGCAGAAATGTCTTACTTCCTTAACC
TTGCTGAAGGTCTAGGTAAAGTTGATGATATCGACCACCTTGGGAATCGTCGTATTCGTGCCGT
TGGTGAATTGCTTGCTAACCAATTCCGTATTGGTCTTGCTCGTATGGAACGTAACGTTCGGGAA
CGTATGTCAGTTCAAGACAACGAAGTGTTGACACCACAACAAATCATCAACATTCGTCCTGTTA

CTGCAGCCGTTAAAGAATTCTTCGGTTCATCTCAATTGTCACAGTTCATGGACCAACACAACCC
ACTTTCTGAGTTGTCTCACAAACGTCGTTTGTCAGCCTTAGGACCTGGTGGTTTGACTCGTGAC
CGTGCTGGTTATGAAGTTCGTGACGTGCACTACACTCACTATGGTCGTATGTGTCCGATTGAAA
CTCCTGAAGGACCTAACATCGGTTTGATCAATAACTTGTCAACATACGGACACCTTAATAAATA
TGGTTTCATCCAAACACCATATCGTAAAGTTGACCGCGCTACAGGTGTGATTACAAACGAAATC
GTTTGGTTGACTGCCGATGAAGAAGATGAATACACAGTAGCACAGGCTAACTCAAAACTTAACG
AAGATGGAACATTTGCTGAAGACATCGTTATGGGACGTCACCAAGGTAATAACCAAGAGTTCCC
AGCAAGCGTTGTTGACTTCGTAGACGTTTCACCTAAACAAGTAGTTGCCGTTGCGACAGCATGT
ATTCCTTTCCTTGAAAACGATGACTCTAACCGTGCCCTTATGGGTGCCAACATGCAACGTCAAG
CGGTGCCATTGATTGATCCACACGCACCATATGTTGGTACTGGTATGGAATATCAAGCAGCCCA
CGACTCAGGTGCTGCAGTTATCGCTAAACACGATGGACGCGTTATCTTCTCTGATGCTGAAAAA
GTTGAAGTTCGTCGCGAAGATGGTTCACTTGATGTTTACCACATTACTAAATTCCGTCGTTCTA
ACTCAGGTACAGCTTATAACCAACATACACTTGTTAAAGTTGGCGATATCGTTGAAAAAGGTGA
CTTCATCGCTGATGGTCCTTCAATGGAAAAAGGTGAAATGGCCCTTGGTCAAAACCCAATCGTC
GCTTACATGACTTGGGATGGTTATAACTATGAAGATGCCATCATCTTGAGTGAACGTCTTGTTA
AAGAAGATGTTTATACATCAGTTCACTTGGAAGAATTTGAATCAGAAACACGTGATACTAAGTT STRF
AGGCCCTGAAGAAATCACTCGCGAAATTCCAAACGTTGGTGAAGAAGCTCTTAAAGACCTTGAC
GAAATGGGTATTATCCGTATCGGTGCTGAAGTTAAAGAAGGTGACATCCTTGTAGGTAAAGTAA
CACCTAAAGGTGAAAAAGACCTTTCTGCTGAAGAGCGCCTTCTTCACGCAATCTTCGGTGATAA
ATCACGTGAAGTTCGTGATACATCACTTCGTGTACCACACGGTGGAGATGGTGTCGTTCGTGAC
GTTAAAATCTTTACACGTGCAAACGGTGATGAATTACAATCAGGTGTTAACATGCTCGTTCGTG
TTTATATCGCACAAAAACGTAAAATCAAAGTCGGAGATAAAATGGCCGGTCGTCACGGTAACAA
AGGGGTTGTTTCTCGTGTTGTTCCAGTTGAAGACATGCCTTATCTTCCAGACGGAACTCCAGTC
GATATCATGTTGAACCCACTTGGGGTGCCATCTCGTATGAACATCGGACAAGTTATGGAGCTTC
ACCTTGGTATGGCTGCTCGTAACCTTGGTATTCACATTGCAACACCAGTCTTTGATGGGGCAAC
TTCTGAAGACCTTTGGGATACAGTTAACGAAGCTGGTATGGCTAGCGACGCTAAGACAGTTCTT
TACGATGGACGTACTGGTGAACCATTTGATAACCGTGTGTCAGTTGGTGTCATGTACATGATTA
AACTTCACCACATGGTTGATGATAAACTTCACGCACGTTCAGTTGGTCCTTACTCACTTGTTAC STRR
GCAACAACCTCTTGGTGGTAAAGCACAATTTGGTGGACAACGTTTCGGTGAAATGGAAGTTTGG
GCTTTGGAAGCTTACGGTGCATCAAATGTTCTTCAAGAAATCTTGACTTACAAATCAGATGATG
TCAACGGTCGTCTTAAAGCTTATGAAGCCATCACTAAAGGTAAACCAATTCCAAAACCAGGTGT
TCCAGAATCATTCCGAGTTCTTGTAAAAGAATTGCAATCACTTGGTCTTGACATGCGCGTGCTT
GATGAAGATGACAATGAAGTAGAACTTCGTGATCTTGATGAAGGTGAAGATGACGATGTTATGC
ACGTTGATGATCTTGAAAAAGCTCGTCAAAAACAAGAAGCAGAAGAAGCGGAAAAAGCAGAAGT
TTCTGCAGAAGAAAACAAATAATAGGAAAGAACATTCAGACATGAGAGAGGCAAGACCTGCTTC
TCTTGGTCAGATTGTTTGATTGAGTCCTATAACGATAAATGATGTCTTACGAATCATGAATTTG
TAAGTCATGACAGTTAGAAAGTAGCGCAGCTATTTCAAAGTCATAAGAAGGTATCATGGTGACG
TAATCGTTACAGCCGGCGTC -3'

SEQ ID n°3: Sequence of the *rpoB* gene of *Abiotrophia defectiva*. This sequence measures 4325 base pairs, has a guanosine plus cytosine content of 47%, and is deposited in GenBank under number AF 535173:

5'-ATATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCCTAAACAACATGTAACGTCACTCCGATG
AGTTGGTTCTGTTGTCTTTTTTTTGCGCTTCAAAGACCGAAAAATGTCATTTGTCAACAATTAT
TAATAATTGTAACCTTAATGTAAAGTGGTGTTCTTAGATTATATTATAGGGGTGAATCGCTTGA
GTCATATCGTGAAATACGGTAAAAAAGCTGAGCGTCGAAGCTATGCGCGTATCGACGAAGTCTT
AGAGTTGCCGAACTTGATTGAAATCCAAACGGATTCCTACAAATGGTTCTTGGATGAAGGGCTA
AAAGTGATGTTCGAGGACATTTCGCCGATTGTCGACCATTCGGAGAACTTGGAACTTCATTTTG
TAGACTATGAGTTCAAGGAAGCTAAGTATAGCTTAGAAGAAGCTCGTAGCCATGACGCTAACTA
CTCAAAACCAATCTATGTAACCTTGCGCCTGTTCAACAAAGAGACAGGTGAAGTCAAAGAACAA
GAAGTCTTCTTCGGGGACTTCCCAATCATGACCGAAATGGGGACCTTCATTATCAACGGGGCGG
AACGGGTTATCGTTTCCCAGTTGGTACGTTCTCCAGGTGTCTACTTCCACGACCGTATGGACAA
GAAAGGCCGCCACAGCTATACTTCTACGGTTATTCCTAACCGTGGGGCTTGGTTGGAATTTGAA
TCAGATGCTAAGGGGATTGCCTACGTCCGCATTGACCGGACCCGGAAGATTCCATTGACTGTCT
TGATGCGTGCCTTAGGTTTTGGTTCAGATGACGAGATTTATGATATCTTCGGCCAATCTGAGCT
CTTAGACTTAACTATCGAGAAGGATGTTCACAAAAACATTCAAGACTCTCGTACGGAAGAAGCC
TTGAAGGACATTTACGAGCGTCTCCGTCCAGGTGAACCTAAGACCGCAGAAAGCTCACGTAACC

TCTTGGTTGCGCGCTTCTTCGACCCACGTCGCTATGACTTAGCACCTGTAGGTCGTTATAAGAT
CAATAAAAAGCTCCACCTCAAGAACCGTTTGGTTGGCTTGACTTTGGCTGAAACCTTGGTTAAC
CCAGAAACAGGCGAAGTGCTCTTTGAAGAAGGAACGGTCTTGGATCAAGAACGTGTTCAAGCCC
TGATTCCATACTTAGAGGCTGGCTTGAATAAGGTAACCCTCTATCCTTCTGAAGATAGTGTGGT
AGCTCAACCAATTGATTTACAAATCATCAAAGTTTATTCACCTAAGAACGCCGAGCAAGTGATT
AACATCATCGGTAACGGGAACATTGAGAAGATTAAGTGCTTGACGCCAGCTGACATTATTGCGT
CAATGAACTACTATCTCTATTTAGACCAAGGAATTGGTGTGACAGATGATATCGACCACTTGGC
TAACCGTCGTATTCGTTCAGTCGGTGAATTATTGCAAAACCAATTCCGTATCGGGCTATCCCGG
ATGGAACGGGTAGTGCGTGAACGTATGTCGCTCCAAGATGTTGCGACCATCACACCGCAACAAT
TGATTAACATTCGTCCAGTAGTGGCGGCTATTAAGGAATTCTTCGGTTCATCCCAGTTGTCACA
ATTCATGGACCAAGTTAACCCACTCGGGGAATTGACCCACAAACGTCGTCTGTCAGCCTTAGGG
CCTGGTGGTTTGACGCGGGACCGTGCCGGCTATGAAGTGCGGGACGTTCACTACTCTCACTACG
GCCGTATGTGTCCAATCGAGACGCCAGAAGGTCCTAACATCGGGTTGATTAACAGCTTGTCTTC
TTATGCCAAGATTAACAAGTATGGTTTTATTGAGACGCCTTACCGTAAAGTGGACAAATCGGTT
ACGCCACACCGTGTCACGACCGAAATTGACTACCTAGCAGCGGACGAGGAAGACTTGTACGTAG
TAGCCCAAGCCAACTCTAAACTCAACGAAGACGGGACCTTCGCCAATGACCTAGTTATGGCGCG
TTTCCGTTCACAAAACATTGAGGTTAACGTTGACCAAGTAGACTACATGGACGTATCGCCAAAA
CAGGTTGTCGCTGTCGCGACTGCTAGCATTCCGTTCTTGGAAAACGACGACTCCAACCGGGGCT
TGATGGGTGCCAACATGCAACGTCAAGCTGTGCCACTTATTAATCCACAATCCCCACTGATTGG
GACTGGGATGGAATATAAGGCAGCACACGACTCTGGGGCTGCGCTCTTATGTAAGCGCGCCGGT
GAAGTGGTTTATGTCGATGCTAACAAGGTGCGCGTGCGCACTCCAGAAGGTGAAGTTGACGAAT
ACCGTTTAACCAAGTTTGCACGTTCTAACGCTGGGACCTGTTACAACCAACGTCCAATCGTAGA
ATTAGGCGACCAAGTTGATGCCTTGGAAATCTTAGCAGATGGTCCATCTATGCAAAATGGGGAG
ATGGCCCTCGGTCAAAACCCACTGGTAGCCTTCATGACTTGGGAAGGGTATAACTATGAGGACG
CGGTTATCATGTCTGAACGTCTGGTCAAAGACGATGTTTATACCTCTATCCACATTGAAGAATA
TGAATCAGAGTCCCGTGAYACYAAGTTAGGCCCTGAAGAAATTACACGCGAAATTCCAAACGTG STRF
TCCGAAGATGCCCTCAAGTACTTAGACAAAGACGGGATTATCTGTATCGGGGCGGAAGTAAAAG
ACGGCGATATCTTAGTTGGTAAGGTAACACCAAAAGGTGTGACCGAGTTGTCTGCGGAAGAACG
CTTGCTCCATGCTATCTTCGGTGAGAAGGCGCGTGAAGTACGTGATACTTCCTTGCGTGTGCCA
CACGGCGGGGGCGGGATTGTCCACGACGTTAAAATCTTTACCCGCGAAGCTGGCGACGAATTGG
CACCAGGTGTCAACAAGCTAGTCCGCGTCTACATCGTACAAAAACGTAAAATCAATGAAGGGGA
TAAGATGGCCGGTCGTCACGGTAACAAAGGGGTTGTCTCCCTTATCATGCCGGAAGAAGATATG
CCATTCTTACCAGATGGTACCCCAGTTGATATCATGTTGAACCCATTAGGGGTTCCATCCCGTA
TGAACATCGGGCAAGTCCTAGAGTTACACTTGGGGATGGCTGCTCGCGAAATGGGCATCAAGAT
TGCAACACCTGTCTTTGACGGTGCTAGTGAAGAAGATGTCTGGGAAACAGTTAAGGAAGCCGGC
TTAGAAGCTGACGCTAAGACTATCTTATATGATGGTCGAACCGGTGAACCATTTGACCGTAAAG
TCTCTGTTGGGGTTATGTACATGATTAAGTTGGCCCACATGGTCGATGACAAGTTGCACGCCCG STRR
TTCAACAGGTCCATACTCTCTGGTTACCCAACAACCATTGGGTGGTAAAGCTCAATTTGGTGGG
CAACGTTTCGGGGAGATGGAGGTTTGGGCCCTA -3'

SEQ ID n° 4: Partial sequence of the *rpoB* gene of *Streptococcus mutans*. This sequence measures 3198 base pairs, has a guanosine plus cytosine content of 42%, and is deposited with GenBank under number AF 535167.

5'-GGACCCTTTTATGACTTCTTGGATACAGGTCTGAAGGAAGTTTTTGAAGATGTGCTTCCAATTT
CCAATTTCACAGACACTATGGAATTAGAGTTTGTGGGTTATGAGTTGAAAGAGCCTAAGTATAC
ATTGGAAGAAGCACGTGCTCATGATGCACATTATTCTGCCCCCATCTTTGTTACTTTCCGTCTC
ATCAATAAAGAAACTGGTGAAATTAAGACACAAGAAGTATTTTTTGGTGATTTTCCCTTGATGA
CTGAAATGGGTACTTTTATTATTAATGGTGCTGAACGTATTATCGTTTCTCAGTTGGTACGTTC
ACCAGGTGTTTATTTTAATGATAAAGTGGATAAAAATGGGAAAATTGGCTATGGTTCAACTGTT
ATCCCTAACCGCGGTGCTTGGCTTGAGCTTGAAACGGACTCTAAGGATATTGCTTATACTCGTA
TTGATCGTACTCGTAAAATTCCTTTTACGACGCTGGTTCGTGCACTCGGTTTTTCCGGGGATGA
TGAGATTATTGATATTTTTGGTGATAGCGAATTGGTTCGTAATACCATTGAAAAAGATATCCAT
AAAAATCCTAATGACTCTCGTACAGATGAAGCTCTCAAGGAANTTATGAACGTCTTCGTCCGGG
TGAACCTAAAACGGCAGATTCNTCACGCAGTCTTCTGATTGCACGTTTCTTTGATGCGCGCCGT
TATGATTAGCAGCTGTTGGCCGCTATAGATAATAAGAAGTTAAACGTCAAAACGGGTCTTTGAA

TCAAGTCATTGGCTGAAAANNAGTAGATCTGAAACAGGCGAAATTCTTGTTGAAAGCTGGGACT
GAAATGACACGCAGTGTAATTGATTCGATTGCAGATTATCTTGATGGAGATCTCAATAAAATTG
TTTATACGCCAAATGAATACGCTGTTTTGACAGAACCTGTTGTTCTTCAAAAATTCAAAGTTAT
GGCTCCAAATGATCCAGACCGCACGGTTACTGTTATTGGTAATGCCAGTCCAAGATGACAAAGT
ACGTCACTTGACACCAGCCGATACGTATTAGCTGAAATGTCTTATTTCCTTAACTTGGCTGAGG
GTNTAGGTAAAGTTGATGATATTGACCATTTAGGCAACCGACGTATTCGTGCTGTTGGTGAATT
GCTTGCTAATCAATTTCGTATTGGTTTGGCACGTATGGAACGCAATGTTCGTGAACGCATGTCC
GTTCAAGATAATGAAGTCTTAACGCCACAACAGATTATTAACATTCGCCCTGTAACAGCGGCAA
TTAAAGAGTTTTTTGGTTCTTCTCAATTGTCACAGTTCATGGACCAACACAATCCACTGTCTGA
ATTGTCTCATAAACGCCGTTTGTCAGCTTTAGGTCCTGGTGGTTTAACACGCGACCGTGCTGGT
TATGAAGTCCGTGATGTGCACTATACGCATTATGGTCGTATGTGTCCAATTGAAACGCCTGAAG
GACCAAATATTGGATTGATTAATAACTTGTCTTCCTATGGTCATCTTAATAAATATGGATTTAT
CCAAACACCATACCGTAAAGTTGACCGTGAGACAGGTAAAGTAACCAATGAAATCGAATGGCTT
ACTGCTGATGAAGAAGATGAATTCACTGTAGCTCAGGCTAACTCAAAACTCAATGAAGATGGAA STRF
GCTTTGCTGAAGAAATCGTCATGGGACGTCATCAAGGGAATAACCAAGAGTTTCCAGCAAGTTC
TGTTGAATATATGGATGTTTCTCCTAAGCAGGTAGTTGCGGTAGCGACAGCATGTATTCCTTTC
CTTGAAAATGATGACTCCAACCGTGCCCTTATGGGAGCTAACATGCAGCGCCAAGCTGTGCCAT
TGATTGATCCTAAAGCACCTTTTGTTGGAACTGGTATGGAATATCAAGCAGCCCATGATTCTGG
AGCCGCTATTATCGCTCAACATAATGGGAAAGTGGTTTATTCCGATGCAGATAAGATTGAAGTT
CGCCGTGAAGATGGCTCACTAGATGTTTATCATGTTACCAAATTCCGTCGTTCTAACTCTGGAA
CTGCCTACAATCAACGTACTCTTGTTAGGGTAGGCGATAGTGTTGAGAAGGGGGACTTTATTGC
AGATGGTCCTTCTATGGAAAAGGGTGAGATGGCTCTTGGACAAAATCCAGTGGTTGCTTACATG
ACTTGGGAGGGTTACAACTTTGAAGATGCTGTTATCATGAGCGAGCGTCTTGTCAAGGATGATG
TTTATACTTCTGTCCATTTAGAAGAATTTGAATCTGAAACTCGTGATACAAAGCTTGGACCTGA
AGAAATTACGCGTGAAATCCCAAATGTTGGTGAAGATGCCCTGAAAGACCTTGATGAAATGGGA
ATTATTCGCATTGGTGCTGAGGTTAAAGAAGGTGATATTCTAGTTGGTAAAGTGACTCCTAAAG
GAGAAAAAGATCTTTCTGCAGAAGAACGCCTCTTGCATGCCATTTTTGGTGACAAATCACGTGA
AGTTCGTGATACTTCTCTTCGTGTACCTCATGGTGGCGACGGTGTTGTTTGTGATGTGAAAATC
TTTACACGTGCTAATGGAGATGAACTTCAATCAGGTGTTAACATGCTGGTTCGTGTTTATATCG
CTCAAAAACGTAAAATCAAGGTCGGAGATAAGATGGCCGGACGTCATGGTAACAAGGGTGTCGT
TTCCCGTATTGTACCAGTGGAAGATATGCCATATCTTCCAGATGGAACACCTGTTGATATCATG
CTTAATCCACTTGGGGTGCCATCACGGATGAACATTGGGCAAGTTATGGAACTCCATCTTGGTA
TGGCTGCTCGTAATTTGGGCATTCATATTGCAACGCCTGTCTTTGACGGAGCAACTTCTGATGA
TCTTTGGGAAACAGTAAAAGAAGCCGGTATGGATTCTGATGCTAAAACTGTTCTTTATGATGGT
CGCACAGGGGAGCCGTTTGATAATCGTGTATCAGTTGGTGTTATGTATATGATTAAACTTCACC STRR
ACATGGTTGATGAYAACCATTTTGTCTATGCAMAGWTCAGTTGGCCCTTAKTCAAYGAWTAMTC
AGASGARTTCCTGCTWGGTGTAAAGGCTNCAATTGTCTTTAGAGGTTAAGGCTGGTGAAATAAC
GGTATGCTGGTATTGATGGCAATGGGCAAGTGAATANTCAACACCGGCCGTCTACANCGTGC-3'

SEQ ID n° 5: Partial sequence of the *rpoB* gene of *Enterococcus faecalis*. This sequence measures 3096 base pairs, has a guanosine plus cytosine content of 42%, and is deposited with GenBank under number AF 535175.

5'-GACCCTTATCAATTGGTTTTTAGATGAGGGACTTCGTGAAATGTTTGAAGACATTTTACCAATT
GATGATTTCCAAGGAAACTTATCCTTAGAATTTGTTGACTATGAATTAAAAGAACCAAAGTACA
CAGTAGAAGAAGCCCGCGCACATGATGCCAACTATTCTGCGCCATTACATGTAACATTACGTTT
AACCAACCGTGAAACAGGTGAAATTAAATCCCAAGAAGTCTTCTTCGGCGATTTCCCATTAATG
ACAGAAATGGGTACCTTCATCATCAACGGGGCAGAACGTGTTATCGTTTCCCAATTAGTTCGTT
CTCCAGGTGTTTACTTCCATGGAAAAGTGGACAAAAACGGCAAAGAAGGTTTTGGCTCAACAGT
CATTCCTAACCGTGGTGCATGGTTAGAAATGGAAACAGATGCGAAAGACATTTCTTATGTTCGG
ATTGACCGCACACGTAAAATTCCTTTAACTGTGTTAGTTCGTGCTTTAGGTTTCGGTTCAGATG
ATACCATCTTCGAAATTTTCGGCGACAGCGAAAGCTTACGCAACACAATTGAAAAAGATTTACA
CAAAAACGCAAGTGATTCTCGTACAGAAGAAGGCTTGAAAGACATTTATGAACGTCTTCGCCCA
GGCGAACCAAAAACAGCAGATAGCTCACGTAGCTTGTTAACTTGCACGTTTCTTTGATCCAAAA
CGTTATGATTTGGCAAACGTTGGTCGCTACAAAGTTAACAAAAAATTAGACTTAAAAACACGTC
TATTAAACTTAACCTTAGCTGAAACGCTAGTTGATCCAGAAACTGGTGTAAATCATTGTCGAAA

AAGGCACAGTTTTAACACACTACATCATGGAAACATTAAGGCRATACATTGACAAACGGCTTAA
ACAGCGTAACTTACTATCCAAGTGAAGATGCGGTAGTAACTGAACCAATGACGATCCAAGTGAT
TCAAGTTCTTTCACCAAAAGATCCTGAACGTATCGTAAATGTGATTGGTAACGGCTATCCAGAC
GACAGCGTAAAAACAGTTCGTCCAGCAGATATCGTTGCTTCAATGAGCTACTTCTTCAACTTAA
TGGAAGATATCGGTAATGTCGATGACATCGACCACTTAGGTAATCGTCGTATCCGTTCAGTAGG
CGAATTATTACAAAACCAATTCCGTATTGGTTTAGCCCGTATGGAACGTGTGGTTCGTGAAAGA
ATGTCTATTCAAGACACAGAAACATTGACACCACAACAATTAATTAACATCCGTCCAGTGGTAG
CAAGTATCAAAGAATTCTTTGGTTCTTCACAGTTATCACAGTTCATGGACCAAACAAACCCATT
AGGTGAGTTAACCCATAAACGTCGTCTATCAGCCTTAGGGCCTGGTGGTTTGACTCGTGATCGT
GCCGGTTATGAAGTTCGTGACGTTCACTACTCTCACTATGGTCGTATGTGTCCAATTGAAACGC
CTGAGGGACCAAATATCGGGTTGATCAATAGCTTATCTAGTTATGCGAAAGTGAATAAATTTGG
TTTCATCGAAACGCCTTATCGCCGTGTTGATCGTGCGACAGGCCGTGTTACTGATCAAGTAGAT
TACTTAACAGCAGACATCGAAGACCATTATATCGTAGCGCAAGCGAACTCACTTTTAAATGAAG
ATGGCACATTTGCCAATGATGTTGTTATGGCGCGTCTACAAAGTGAAAACTTAGAAGTTGCCGT
AGACAAAGTTGACTACATGGACGTTTCACCAAAACAAGTAGTCGCAGTCGCAACAGCATGTATT
CCTTTCTTAGAAAACGATGACTCCAACCGTGCCTTGATGGGTGCCAACATGCAGCGTCAAGCGG
TGCCGTTAATTCAACCACGCTCTCCGTGGGTAGGTACAGGTATGGAATATAAATCAGCCCATGA
CTCAGGTGCTGCTTTACTATGTAAACATGACGGTGTCGTAGAATTCGTCGATGCAAAAGAAATT STRF
CGCGTTCGTCGCGACAATGGCGCATTAGACAAATATATGGTTACAAAATTCCGTCGTTCTAACT
CAGGAACAAGCTACAACCAACGCCCAATTGTTCACTTAGGTGAAAAGTTGAAAAGGCGATACTT
TACCGGATGGACCTTCTATGGAAGAAGCGAAATGGCTTTATGGCAAAACGTCTTAGTTGCCTTC
ATGACATGGGAAGGTTACAACTACGAGGATGCCATTATCATGAGCCGTCGTTTAGTTAAAGACG
ATGTCTACACTTCTGTGCATATTGAAGAATATGAATCAGAAGCACGTGATACAAAATTAGGACC
TGAAGAAATTACCCGTGAAATTCCAAACGTTGGGGAAGACGCGTTGAAAGACTTAGACGAAATG
GGGATTATCCGCATTGGTGCTGAAGTTCAAGATGGCGACTTACTAGTTGGGAAAGTCACACCTA
AAGGGGTCACAGAATTATCTGCAGAAGAACGTTTATTACACGCAATCTTCGGGGAAAAAGCCCG
CGAAGTTCGTGATACGTCTCTCCGTGTACCTCACGGTGGCGGCGGTATCGTTCATGATGTGAAA
ATCTTTACTCGTGAAGCTGGCGATGAATTATCACCAGGTGTCAACATGTTAGTTCGTGTCTATA
TCGTTCAAAAACGTAAAATTCACGAAGGAGATAAAATGGCGGGACGTCACGGAAATAAAGGGGT
TGTTTCCCGTATTATGCCGGAAGAAGATATGCCATTCTTACCTGACGGAACACCTGTTGATATC
ATGTTGAACCCATTAGGGGTACCTTCTCGTATGAATATCGGACAAGTACTTGAATTACACTTAG
GTATGGCTGCTCGCCAATTAGGTATTCACGTCGCAACACCTGTTTTCGATGGGGCAACCGATGA
AGACGTTTGGGAAACTGTTCGTGAAGCTGGTATGGCTAGCGATGCTAAAACAGTTCTTTACGAT
GGACGTACAGGTGAACCATTTGATAACCGTATTTCCGTTGGTGTCATGTATATGATTAAATTAG
CCCACATGGTTGATGACAAATTGCATGCTCGTTCAATCGGACCTTACTCTCTTGTTACGCAACA STRR
ACCGTTGGGTGTAAAGCTCAATTC-3'

In the preceding sequences, the K nucleotide designates T or G, the M nucleotide designates A or C, the R nucleotide designates A or G, the W nucleotide designates A or T, the Y nucleotide designates C or T and the N nucleotide designates A, T, C or G.

Example 2: Partial sequencing of the *rpoB* gene of 28 species of genus *Streptococcus* and related genera.

From the alignment of the complete sequences of the *rpoB* gene in *Streptococcus spp.* and *Abiotrophia defectiva* in example 1 and those known in GenBank (*Streptococcus pneumoniae* AE008542 and *Streptococcus pyogenes* AE006480) a set of primers was chosen for the amplification and sequencing of a 709 to

740 bp fragment of this gene in 28 type strains of these bacterial genera. The sequences of these primers were:

- SEQ ID n° 6: 5'- AARYTIGMCCTGAAGAAAT-3'
- SEQ ID n° 7: 5'- TGIARTTTRTCATCAACCATGTG-3'

Sequence SEQ ID n° 7 was used as 3' primer and therefore represents the complementary reverse sequence of the direct strand represented in preceding sequences SEQ ID n° 1 to 5.

These primers are incorporated with the DNA extracted from the bacteria during PCR under the following conditions: denaturing at 95°C for 1 min followed by 35 cycles comprising a denaturing step at 94°C for 10 sec, a hybridisation step at 52°C for 10 sec and an elongation step at 72°C for 30 sec.

The amplified products are sequenced with the same primers SEQ ID n° 6 and SEQ ID n° 7 under the following conditions: denaturing at 95°C for 1 min followed by 30 cycles comprising a denaturing step at 95°C for 30 sec, a hybridisation step at 52°C for 30 sec and a hybridisation step at 62°C for 1 min. The sequencing products are analysed on a ABI PRISM 3100 sequencer.

The inventors determined the position of these two primers SEQ ID n° 6 and SEQ ID n° 7, so as to observe the following criteria:

1- sequence flanked by these two primers specific to the species of the bacterium. This condition is verified after alignment of the fragments of around 720 bp with all the sequences of the *rpoB* bacterial genes available in computerized data banks,

2- search for the shortest possible identification region to achieve the best possible increase in the sensitivity of molecular detection,

3- primer length of 18 to 22 bp,

4- sequence of primers showing a close melting temperature,

5- sequence of primers not enabling auto-hybridisation or complementarity

The obtained *rpoB* gene fragments of the bacterial species of genus *Streptococcus* and said related genera have approximately 720 (709 to 732) base pairs and their sequence is specific to each species of this genus therefore permitting molecular identification of the bacteria of the 28 species tested, i.e.:

SEQ ID n° 8 : partial sequence of the *rpoB* gene in *Streptococcus suis* CIP 1032 17^T measuring 709 base pairs:

5' – CGCGAAATTCCAAACGTTGGTGAAGATGCCCTTCGCAACTTGGACGAAA
TGGGGATTATCCGTATTGGTGCCGAAGTTAAAGAGGGCGACATTCTTGTTGG
TAAAGTCACACCAAAAGGTGAAAAAGATCTTTCTGCTGAAGAGCGTCTCTTGC
ACGCAATCTTCGGTGACAAGTCACGTGAAGTACGTGATACCTCTCTTCGTGTA
CCTCACGGTGCCGATGGTGTCGTTCGTGATGTGAAAATCTTTACTCGTGCCAA
CGGTGATGAATTGCAATCAGGTGTTAACATGTTGGTTCGTGTTTACATCGCTC
AAAAACGTAAGATCAAGGTCGGAGATAAGATGGCCGGTCGTCACGGTAACAA
GGGTGTCGTTTCACGTATTGTACCTGTTGAGGATATGCCATATCTTCCAGATG
GAACACCAGTTGACATCATGTTGAACCCACTCGGGGTGCCATCACGTATGAAC
ATCGGTCAGGTTATGGAACTTCACTTGGGTATGGCGGCTCGCAACTTGGGCA
TCCATATCGCAACACCAGTTTTCGATGGTGCAAGTTCAGAAGACCTCTGGTCA
ACTGTTAAAGAAGCAGGTATGGACTCAGATGCCAAGACCATTCTTTACGATGG
ACGTACAGGTGAACCATTTGACAACCGTGTATCTGTTGGTGTCATGTACATGA
TCAAGCTTCACCACATGGTTGATGACA – 3'

SEQ ID n° 9: partial sequence of the *rpoB gene* in *Streptococcus sanguinis* CIP 55.128^T measuring 725 base pairs:

5'- TGTCATCAACCATGTGGTGAGCTTAATCATGTACATGACACCGACAGATA
CACGGTTGTCAAACGGCTCACCGGTACGTCCATCGTAAAGAATAGTCTTGGCA
TCGCTATCCATACCAGCTTCACGGACAGTATCCCAGAGGTCTTCTGAGCTTGC
TCCATCAAAGACCGGTGTCGCAATATGGATGCCCAAGTTACGTGCTGCCATAC
CAAGGTGAAGCTCCATAACCTGACCAATGTTCATACGTGATGGTACCCCGAGT
GGGTTCAGCATGATATCAACTGGTGTTCCGTCTGGCAAATAAGGCATGTCTTC
CACAGGAACGATACGGGATACAACCCCCTTGTTTCCGTGACGACCAGCCATCT
TATCTCCGACCTTGATCTTACGTTTTTGAGCGATGTAGACACGAACCAACATAT
TAACGCCAGATTGCAACTCATCACCATTAGCACGGGTAAAGATCTTCACGTCA
CGAACCACTCCATCAGCACCGTGCGGCACACGCAGAGAGGTATCACGGACTTC
ACGAGACTTGTCTCCGAAGATAGCGTGCAAGAGGCGCTCTTCAGCAGAAAGA
TCTTTTTCACCCTTAGGGGTAACTTTACCTACAAGGATATCGCCTTCCTTGACT
TCCGCCCCGATGCGGATAATACCCATTTCGTCCAAATTGCGTAGGGCATCTTC
CCCTACGTTTGGAATTTCGCGGGTAATTCTTCAGGTCA – 3'

SEQ ID n°10: partial sequence of the *rpoB* gene in *Streptococcus salivarius* CIP 102503^T measuring 728 base pairs:

5'- TTGTCATCAACCATGTGTGAAGTTTGATCATGTACATGACACCAACTGAT
ACACGGTTATCAAATGGTTCACCTGTACGTCCATCGTAAAGGATTGTCTTAGC
ATCACTATCCATACCTGCTTCACGAACAGTATCCCAGAGGTCTTCTGAGCTTGC
CCCGTCAAAGACTGGTGTTGCGATGTGGATACCCAAGTTACGAGCAGCCATA
CCAAGGTGAAGTTCCATAACCTGACCGATGTTCATACGTGATGGCACCCCAAG
AGGGTTCAACATGATATCAACTGGTGTACCGTCTGGAAGGTAAGGCATGTCT
TCAACAGGAACAATACGAGAAACAACCCCTTTGTTACCGTGACGACCGGCCAT
CTTATCTCCGACCTTAATCTTACGTTTTTGAGCGATGTAAACACGAACAAGCAT
GTTAACACCTGATTGCAATTCATCACCGTTTGCACGTGTGAAGATTTTAACATC
ACGAACGACACCATCACCACCGTGAGGTACACGGAGTGAGGTATCACGTACT
TCACGAGATTTATCACCAAAGATAGCATGGAGAAGACGTTCTTCAGCAGAAA
GGTCTTTTTCACCCTTAGGTGTTACCTTACCAACAAGAATGTCACCTTCTTTAA
CCTCAGCACCGATACGGATAATACCCATTTCGTCAAGGTCTTTGAGAGCTTCTT
CACCAACGTTTGGCAATTCACGTGTAATTTCTTCAGGTCCA – 3'

SEQ ID n°11: partial sequence of the *rpoB* gene in *Streptococcus pyogenes* CIP 56.41^T measuring 725 base pairs:

5'- TGTCATCAACCATGTGGTGAAGTTTGATCATATACATGACACCAACGGAT
ACACGGTTGTCAAATGGTTCACCGGTGCGACCATCATAAAGGACCGTCTTAGC
ATCGCTATCCATACCAGCTTCACGAACAGTGTCCCAAAGGTCTTCTGATGAAG
CCCCGTCAAAGACAGGTGTTGCAATGTGAATACCAAGATTACGAGCAGCCATA
CCAAGGTGAAGTTCCATAACCTGACCAATATTCATCCGTGATGGCACCCCAAG
AGGGTTCAACATGATGTCAACTGGTGTTCCGTCTGGAAGGTATGGCATGTCT
TCAACTGGTACAATACGTGAAACGACACCCTTGTTTCCGTGACGACCGGCCAT
TTTATCTCCGACCTTGATTTTACGTTTTTGAGCGATGTAAACACGCACAAGCAT
ATTAACACCTGATTGCAATTCATCGCCGTTAGCGCGTGTAAAGATTTTCACATC
ACGAACGATACCATCACCACCGTGAGGGACACGAAGTGAGGTATCACGCACT
TCACGCGATTTATCCCCAAAGATGGCGTGAAGTAAACGTTCTTCAGCAGAAAG
GTCTTTTTCACCTTTAGGTGTGACTTTACCTACTAAGATGTCGCCTTCTTTAAC
CTCAGCACCGATACGGATAATGCCCATTTCGTCAAGGTCTTTGAGGGCTTCTT
CACCAACATTTGGGATTTCCGAGTGATTCTTCAGGGCA – 3'

SEQ ID n°12: partial sequence of the *rpoB* gene in *Streptococcus pneumoniae* CIP 102911^T measuring 724 base pairs:

5' – CAACCATGTGGTGGAGTTTGATCATGTACATGACTCCGACAGAAAACACG
GTTATCAAACGGTTCACCAGTACGTCCATCGTAAAGGATCGTTTTGGCATCGC
TATCCATACCTGCTTCTTTAACAGTTGACCAAAGATCTTCAGAACTTGCTCCAT
CAAAGACTGGTGTCGCGATGTGAATACCAAGAGTACGAGCTGCCATACCAAG
GTGAAGCTCCATAACCTGACCGATATTCATACGTGATGGTACCCCAAGTGGGT
TCAACATGATGTCGACTGGAGTTCCGTCTGGAAGGTAAGGCATGTCTTCTACA
GGAACGATACGAGAGACAACCCCTTTGTTTCCGTGACGTCCGGCCATTTTATC
TCCGACCTTAATCTTACGTTTTTGAGCGATGTAAACACGAACCAACATGTTAAC
ACCTGATTGCAACTCATCTCCATTTACACGTGTAAAGATCTTAACATCACGAAC
GACACCATCGGCACCGTGTGGTACACGAAGAGAAGTATCACGCACTTCACGA
GACTTGTCTCCAAAGATAGCGTGCAAGAGACGTTCTTCAGCTGAAAGATCTTT
CTCACCCTTAGGTGTTACTTTACCTACAAGAATATCACCTTCTTTAACCTCAGCA
CCAATACGGATAATCCCATTTCGTCAAGGTCTTTGAGGGCATCTTCACCAACG
TTTTGGAATTTCGCGAGTGATTTCTTCAGGTCCAA – 3'

SEQ ID n°13: partial sequence of the *rpoB* gene in *Streptococcus oralis* CIP 102922^T measuring 694 base pairs:

5'-

ACTCGTGAAATTCCAAACGTTGGTGAAGATGCCCTTAAAGACCTTGACGAAAT
GGGTATTATCCGTATTGGTGCTGAGGTTAAAGAAGGAGATATCCTTGTAGGT
AAAGTCACACCTAAGGGTGAAAAAGACCTTTCTGCTGAAGAACGTCTCTTGCA
CGCTATCTTCGGAGACAAGTCTCGTGAAGTGCGTGATACTTCTCTTCGAGTAC
CTCACGGTGCCGATGGTGTCGTTCGTGATGTTAAGATCTTTACACGTGCAAAT
GGTGATGAGTTGCAATCTGGTGTGAATATGCTGGTTCGTGTCTACATCGCTCA
AAAACGTAAGATCAAGTCGGAGATAAGATGGCCGGACGTCACGGAAACAAAG
GGGTTGTCTCTCGTATCGTTCCTGTAGAAGACATGCCTTACCTTCCAGATGGA
ACTCCAGTCGATATCATGTTGAACCCACTTGGGGTGCCATCACGTATGAATAT
CGGTCAGGTTATGGAACTCCACCTTGGTATGGCAGCCCGTACTCTTGGTATCC
ACATCGCAACACCAGTCTTTGACGGAGCAAGTTCGGAAGACCTTTGGGACACT
GTTAAAGAAGCAGGTATGGATAGCGATGCCAAAACAATCCTTTACGATGGAC
GTACAGGTGAGCCGTTTGACAACCGTGTATCAGTTGGTGTCATGTACATGATC
AAACTCCA– 3'

SEQ ID n°14: partial sequence of the *rpoB* gene in *Streptococcus mutans* CIP 103220^T measuring 728 base pairs:

5' – TGTCATCAACCATGTGGTGAAGTTTAATCATATACATAACACCAACTGATA
CACGATTATCAAACGGCTCCCCTGTGCGACCATCATAAAGAACAGTTTTAGCA
TCAGAATCCATACCGGCTTCTTTTACTGTTTCCCAAAGATCATCAGAAGTTGCT
CCGTCAAAGACAGGCGTTGCAATATGAATGCCCAAATTACGAGCAGCCATACC
AAGATGGAGTTCCATAACTTGCCCAATGTTCATCCGTGATGGCACCCCAAGTG
GATTAAGCATGATATCAACAGGTGTTCCATCTGGAAGATATGGCATATCTTCC
ACTGGTACAATACGGGAAACGACACCCTTGTTACCATGACGTCCGGCCATCTT
ATCTCCGACCTTGATTTTACGTTTTTGAGCGATATAAACACGAACCAGCATGTT
AACACCTGATTGAAGTTCATCTCCATTAGCACGTGTAAAGATTTTCACATCACA
AACAACACCGTCGCCACCATGAGGTACACGAAGAGAAGTATCACGAACTTCAC
GTGATTTGTCACCAAAAATGGCATGCAAGAGGCGTTCTTCTGCAGAAAGATCT
TTTTCTCCTTTAGGAGTCACTTTACCAACTAGAATATCACCTCTTTAACCTCAG
CACCAATGCGAATAATTCCCATTTCATCAAGGTCTTTCAGGGCATCTTCACCAA
CATTTGGGATTTCACGCGTAATTTCTTCAGGTCCA – 3'

SEQ ID n°15: partial sequence of the *rpoB* gene in *Streptococcus mitis* CIP 103335^T measuring 730 base pairs:

5'- TGTCATCAACCATGTGGTGGAGTTTGATCATGTAACATGACTCCGACAGA
AAACACGGTTATCAAATGGTTCACCTGTACGTCCATCGTAAAGGATTGTTTTG
GCATCGCTATCCATACCAGCTTCTTTAACAGTTGACCAAAGATCTTCAGAACTT
GCTCCGTCAAAGACTGGTGTTGCGATGTGAATACCAAGAGTACGAGCTGCCA
TCCCAAGGTGGAGTTCCATAACCTGACCGATATTCATACGTGATGGCACCCCA
AGTGGGTTCAACATGATATCGACTGGAGTTCCATCTGGAAGGTAAGGCATAT
CTTCTACAGGAACGATACGAGAGACAACCCCTTTATTTCCGTGACGTCCGGCC
ATCTTATCTCCGACCTTGATCTTACGTTTTTGAGCGATGTAGACGCGAACCAG
CATGTTGACACCTGATTGCAATTCATCTCCATTTGCACGTGTAAAGATCTTAAC
ATCACGAACCACACCATCAGCTCCGTGTGGTACACGAAGAGAAGTGTCACGTA
CTTCACGAGATTTATCTCCGAAGATAGCGTGCAAGAGCCGTTCTTCAGCTGAA
AGGTCTTTCTCACCCTTAGGTGTTACTTTACCTACAAGGATATCCCCTTCTTTA
ACCTCAGCACCGATACGGATAATACCCATTTCGTCAAGATCTTTAAGGGCATC
TTCCCCAACGTTTGGGATTTCACGAGTAATTTCTTCAGGTCCA – 3'

SEQ ID n°16: partial sequence of the *rpoB* gene in *Streptococcus equinus* CIP 102504^T measuring 697 base pairs:

5'-
CACTCGCGAAATTCCAAACGTTGGTGAAGAAGCTCTTAAAGACCTTGACGAAA
TGGGTATTATCCGTATCGGTGCTGAAGTTAAAGAAGGTGACATCCTTGTAGG
TAAAGTAACACCTAAAGGTGAAAAAGACCTTTCTGCTGAAGAGCGCCTTCTTC
ACGCAATCTTCGGTGATAAATCACGTGAAGTTCGTGATACATCACTTCGTGTA
CCACACGGTGGAGATGGTGTCGTTCGTGACGTTAAAATCTTTACACGTGCAAA
CGGTGATGAATTACAATCAGGTGTTAACATGCTCGTTCGTGTTTATATCGCAC
AAAAACGTAAAATCAAAGTCGGAGATAAAATGGCCGGTCGTCACGGTAACAA
AGGGGTTGTTTCTCGTGTTGTTCCAGTTGAAGACATGCCTTATCTTCCAGACG
GAACTCCAGTCGATATCATGTTGAACCCACTTGGGGTGCCATCTCGTATGAAC
ATCGGACAAGTTATGGAGCTTCACCTTGGTATGGCTGCTCGTAACCTTGGTAT
TCACATTGCAACACCAGTCTTTGATGGGGCAACTTCTGAAGACCTTTGGGATA
CAGTTAACGAAGCTGGTATGGCTAGCGACGCTAAGACAGTTCTTTACGATGG
ACGTACTGGTGAACCATTTGATAACCGTGTGTCAGTTGGTGTCATGTACATGA
TTAAACTTCAC– 3'

SEQ ID n°17: partial sequence of the *rpoB* gene in *Streptococcus constellatus* CIP 103247^T measuring 731 base pairs:

5'- AGTTGTCATCAACCATGTGTGCAATTTAATCATATACATGACACCGACAGA
TACACGGTTGTCAAACGGCTCGCCCGTACGACCATCATAAAGAATCGTCTTGG
CATCGCTATCCATGCCTGCTTCACGAACAGTATCCCAAAGGTCATCTGAGCTT
GCTCCGTCAAATACTGGCGTTGCTATGTGGATACCAAGGTTGCGAGCAGCCA
TACCAAGGTGAAGCTCCATAACCTGTCCGATATTCATACGTGATGGCACCCCA
AGTGGGTTCAACATGATGTCTACTGGTGTTCCGTCTGGAAGATAAGGCATAT
CCTCAACTGGAACGATACGGGAAACAACCCCTTTATTTCCGTGGCGTCCGGCC
ATCTTATCCCCAACGCGGATCTTTCGTTTTTGAGCAATGTAAACACGCACCAAC
ATGTTGACACCAGATTGCAATTCATCACCGTTCGCACGAGTAAAGATTTTCAC
ATCACGGACAACCCCAGCACCACCATGTGGTACACGAAGAGATGTGTCACGTA
CTTCACGAGATTTATCACCGAAAATTGCATGAAGCAGGCGTTCTTCAGCGGAT
AAGTCTTTTTCACCTTTCGGCGTTACTTTACCGACAAGAATGTCGCCCTCTTTC
ACCTCAGCACCAATGCGGATAATTCCCATTTCGTCAAGGTCTCTTAGCGCATCT
TCCCCAACGTTTGGAATTTCGCGCGTAATTTCTTCAGGTCCAA – 3'

SEQ ID n°18: partial sequence of the *rpoB* gene in *Streptococcus anginosus* CIP 102921^T measuring 697 base pairs:

5' –
CACGCGCGAAATTCCAAACGTCGGTGAAGATGCTTTGAGAGACCTTGACGAA
ACGGGAATTATCCGCATTGGTGCTGAGGTAAAAGAAGGCGACATTCTTGTCG
GTAAAGTAACACCGAAAGGTGAAAAAGACTTATCTGCTGAAGAACGCCTGCT
TCATGCAATTTTCGGTGATAAATCTCGTGAAGTACGTGATACTTCCCTTCGTGT
ACCACATGGTGGTGCAGGGGTTGTCCGTGATGTGAAAATCTTTACTCGTGCG
AACGGTGATGAATTGCAATCTGGTGTCAACATGTTGGTACGTGTTTACATCGC
TCAAAAACGGAAAATCCGTGTTGGGGATAAGATGGCTGGACGTCACGGAAAC
AAAGGGGTTGTTTCCCGCATTGTTCCAGTTGAGGATATGCCGTATCTTCCAGA
TGGAACACCAGTTGATATTATGTTGAACCCACTTGGGGTGCCATCTCGTATGA
ATATTGGTCAAGTTATGGAGCTTCACCTCGGTATGGCTGCTCGCAACCTTGGC
ATTCACATTGCAACACCAGTATTTGACGGGGCTAGCTCAGATGATCTTTGGGA
AACCGTTCGTGAAGCTGGCATGGATAGCGATGCTAAGACAATCCTTTATGAT
GGCCGTACTGGTGAGCCATTTGATAATCGTGTATCCGTTGGTGTCATGTACAT
GATCAAACTCCAC– 3'

SEQ ID n°19: partial sequence of the *rpoB* gene in *Streptococcus dysgalactiae* CIP 102914^T measuring 728 base pairs:

5' – TGTCATCAACCATGTGGTGGAGTTTAATCATGTACATGACACCAACGGAT
ACACGGTTGTCAAATGGTTCGCCAGTACGTCCATCATAAAGGACCGTCTTAGC
ATCGCTATCCATACCAGCTTCACGAACAGTGTCCCAAAGGTCTTCTGATGAAG
CCCCGTCAAAGACAGGTGTTGCAATGTGAATACCAAGATTACGAGCAGCCATA
CCAAGGTGAAGTTCCATAACCTGACCAATGTTCATCCGTGATGGCACCCCAAG
AGGGTTCAACATGATGTCAACTGGTGTTCCATCTGGAAGGTATGGCATGTCTT
CAACTGGTACAATACGTGAAACGACACCCTTGTTTCCGTGACGACCAGCCATT
TTATCTCCGACTTTGATCTTACGTTTTTGAGCAATGTAAACACGCACAAGCATA
TTAACACCTGATTGCAATTCATCGCCGTTAGCGCGTGTAAAGATTTTCACATCA
CGAACGATACCATCACCACCGTGAGGTACACGAAGGGACGTATCACGAACTTC
ACGTGATTTATCTCCAAAGATGGCATGCAAGAGACGCTCTTCAGCAGAAAGGT
CTTTTTCACCTTTAGGTGTGACTTTACCTACTAAGATGTCGCCTTCTTTAACCTC
AGCACCGATACGGATAATTCCCATTTCGTCAAGGTCTTTGAGCGCTTCTTCACC
AACGTTTGGAATTTCGCGGGTGATTTCTTCAGGTCAA – 3'

SEQ ID n°20: partial sequence of the *rpoB* gene in *Streptococcus bovis* CIP 102302^T measuring 728 base pairs:

5' – TGTCATCAACCATGTGGTGAAGTTTGATCATGTACATGATACCAACAGAG
ACACGATTATCAAATGGTTCACCTGTACGACCGTCATAAAGAACTGTCTTAGC
GTCGCTATCCATACCAGCTTCACGAACAGTATCCCAAAGGTCTTCTGAAGTTG
CCCCGTCAAAGACTGGAGTTGCAATGTGAATACCGAGGTTACGAGCTGCCAT
ACCAAGGTGAAGTTCCATAACTTGTCCGATATTCATACGAGATGGCACCCCAA
GAGGGTTCAACATGATATCAACTGGAGTTCCGTCTGGAAGATATGGCATGTC
TTCAACAGGAACGATACGAGAAACAACCCCTTTGTTTCCGTGACGACCGGCCA
TTTTATCTCCGACTTTGATTTTACGTTTTTGTGCAATGTAAACACGAACGAGCA
TGTTGACACCTGATTGCAATTCATCACCGTTAGCACGTGTGAAGATTTTAACA
TCACGAACAACACCGTCTCCACCGTGTGGCACACGAAGTGATGTATCACGTAC
TTCACGAGATTTATCACCGAAGATTGCGTGAAGAAGGCGTTCTTCAGCAGAAA
GGTCTTTTTCACCTTTAGGTGTTACTTTACCTACAAGGATATCACCTTCTTTAA
CTTCAGCACCGATACGGATAATACCCATTTCGTCAAGGTCTTTAAGAGCTTCTT
CACCAACGTTTGGAATTTCGCGAGTGATTTCTTCAGGTCAA – 3'

SEQID n°21: partial sequence of the *rpoB* gene in *Streptococcus acidominimus* CIP 82.4^T measuring 728 base pairs:

5'- TTGTCATCAACCATGTGGTGGAGCTTAATCATGTACATGACACCAACAG
ACACACGGTTATCAAATGGTTCACCAGTACGACCATCATAAAGAATCGTTTTA
GCATCGCTGTCCATTCCTGCCTCTTTAACAGTTGACCAGAGATCCTCTGAGCTC
GCACCATCGAAAACCGGTGTTGCGATATGGATACCCAAGTTACGAGCAGCCAT
ACCCAAGTGCAGTTCCATAACCTGACCAATATTCATACGAGATGGCACCCCAA
GTGGGTTCAACATGATGTCAACTGGTGTTCCATCTGGAAGATATGGCATGTCT
TCAACTGGTACAATACGAGAAACGACACCCTTGTTACCGTGACGACCGGCCAT
CTTATCTCCGACCTTAATCTTGCGTTTTTGAGCGATATACACACGTACCAGCAT
ATTAACACCAGACTGTAGCTCATCACCATTAGCACGCGTAAAGATTTTCACATC
ACGAACAACACCATCTGCACCGTGTGGCACACGTAGAGAGGTATCACGTACTT

CACGTGATTTGTCACCGAAGATAGCATGCAAGAGACGCTCCTCAGCAGAAAG
ATCTTTTTCACCTTTTGGTGTCACCTTACCAACAAGAATATCGCCTTCTTTAACT
TCTGCACCGATACGGATAATACCCATTTCGTCAAGGTCTTTGAGGGCTTCTTC
ACCAACGTTTGGAATTTCACGAGTAATTTCTTCAGGTCA – 3'

SEQ ID n°22: partial sequence of the *rpoB* gene in *Streptococcus agalactiae* CIP 103227^T measuring 733 base pairs:

5' – TGAGTTGTCATCAACCATGTGGTGAAGTTTGATCATGTACATGACACCAA
CTGACACACGGTTATCGAATGGTTCACCAGTACGACCATCATAAAGAACAGTC
TTAGCATCTGAATCCATACCTGCTTCTTGAACAGTTTCCCAAAGGTCTTCTGAA
GAAGCCCCATCAAAGACTGGCGTTGCAATATGAATACCTAAATTACGAGCAGC
CATACCTAAATGAAGCTCCATAACTTGTCCGATATTCATACGTGATGGCACCCC
AAGTGGGTTCAACATGATATCAACTGGCGTTCCATCTGGTAAGTAAGGCATAT
CTTCAACAGGAACAATACGTGAGACGACACCTTTGTTTCCGTGACGACCGGCC
ATCTTATCACCGACTTTGATTTTACGTTTTTGAGCGATATAAACGCGGACAAG
CATATTAACACCTGATTGCAATTCATCACCATTTGCACGAGTAAAGATTTTAAC
GTCACGAACTACTCCATCGCCACCGTGAGGTACACGTAGTGAAGTATCACGAA
CTTCACGTGATTTATCACCAAAAATGGCATGCAAGAGACGTTCTTCAGCAGAT
AAGTCCTTTTCACCCTTAGGTGTTACCTTACCAACAAGAATGTCACCTTCTTTT
ACCTCAGCACCAATGCGGATAATTCCCATTTCATCGAGATCACGTAGTGAATC
TTCACCAACATTTTGGATTTCACGAGTAATTTCTTCAGGTCCA – 3'

SEQ ID n°23: partial sequence of the *rpoB* gene in *Streptococcus difficilis* CIP 103768^T measuring 714 base pairs:

5'- TTGTCATCAACCATGTGGTGAAGTTTGATCATGTACATGACACCAACTGAC
ACACGGTTATCGAATGGTTCACCAGTATGACCATCATAAAGAACAGTCTTAGCAT
CTGAATCCATACCTGCTTCTTGAACAGTTTCCCAAAGGTCTTCTGAAGAAGCCCC
ATCAAAGACTGGCGTTGCAATATGAATACCTAAATTACGAGCAGCCATACCTAAA
TGAAGCTCCATAACTTGTCCGATATTCATACGTGATGGCACCCCAAGTGGGTTCA
ACATGATATCAACTGGCGTTCCATCTGGTAAATAAGGCATATCTTCAACAGGAAC
AATACGTGAGACGACACCTTTGTTTCCGTGACGACCGGCCATCTTATCACCGACT
TTGATTTTACGTTTTTGAGCGATATAAACGCGGACAAGCATATTAACACCTGATT
GCAATTCATCACCATTTGCACGAGTAAAGATTTTAACGTCACGAACTACTCCATC
GCCACCGTGAGGTACACGTAGTGAAGTATCACGAACTTCACGTGATTTATCACCA
AAAATGGCATGCAAGAGACGTTCTTCAGCAGATAAGTCCTTTTCACCCTTAGGCG
TTACCTTACCAACAAGAATGTCACCTTCTTTTACCTCAGCACCAATGCGGATAATT
CCCATTTCATCGAGATCACGTAGTGAATCTTCACCAACATTTGGAATTTCACGAG
TA – 3'

SEQ ID n°24: partial sequence of the *rpoB* gene in *Streptococcus intermedius* CIP 103248^T measuring 728 base pairs:

5'- TGTCATCAACCATGTGGTGAAGCTTAATCATGTACATGACACCAACGGAC
ACACGGTTATCAAACGGTTCGCCAGTACGTCCATCATAAAGGATTGTCTTAGC
ATCGCTATCCATACCTGCTTCACGAACGGTTTCCCAAAGATCATCTGAGCTAGC
TCCGTCAAAGACTGGCGTTGCAATGTGGATACCAAGTTGCGAGCAGCCATAC
CGAGGTGCAATTCCATAACTTGTCCGATATTCATACGTGACGGCACCCCAAGA
GGATTCAACATGATATCAACTGGTGTCCCGTCTGGAAGATACGGCATATCCTC
AACTGGAACAATGCGGGAAACAACCCCTTTGTTTCCGTGGCGTCCGGCCATCT
TATCTCCAACGCGGATTTTCCGTTTTTGAGCGATATAAACACGTACCAACATGT
TGACACCGGATTGCAATTCATCACCGTTCGCACGAGTAAAGATTTTTACATCAC
GGACAACACCTGCACCACCGTGTGGTACACGAAGGGAGGTATCACGCACTTC
ACGAGACTTATCACCAAAAATTGCATGAAGCAGGCGTTCTTCAGCGGATAAAT
CTTTTTCACCTTTCGGCGTTACTTTACCGACAAGAATGTCGCCTTCTTTTACCTC
AGCACCAATGCGGATAATTCCCATCTCGTCAAGGTCTCTCAAAGCATCTTCCCC
GACGTTTGGAATTTCGCGCGTGATTTCTTCAGGTCCA – 3'

SEQ ID n°25: partial sequence of the *rpoB* gene in *Streptococcus equi* CIP 102910^T measuring 728 base pairs:

5'- TGTCATCAACCATGTGGTGAAGCTTAATCATATACATGACACCAACTGAC
ACACGATTATCAAACGGCTCACCAGTACGGCCATCATAAAGAACAGTCTTAGC
ATCGCTATCCATACCTGCTTCACGAACAGTTTCCCAAAGGTCCTCAGACGTAGC
TCCGTCAAAGACCGGTGTTGCGATATGGATACCCAAATTACGAGCAGCCATAC
CTAGGTGAAGCTCCATAACCTGTCCAATGTTCATACGAGACGGCACCCCAAGA
GGGTTCAGCATGATGTCAACAGGGGTTCCGTCTGGCAGATATGGCATATCCT
CAACCGGTACAATACGTGAGACGACACCCTTGTTACCATGACGCCCGGCCATT
TTATCTCCGACCTTGATTTTACGCTTTTGAGCAATGTAAACACGCACCAGCATA
TTAACACCTGATTGAAGCTCATCACCATTTGCGCGTGTAAAGATCTTCACATCA
CGTACAATCCCGTCACCACCATGAGGAACACGTAACGAGGTATCACGAACCTC
ACGTGATTTATCACCAAAGATAGCATGCAGGAGACGTTCTTCAGCAGAAAGG
TCTTTTTCACCCTTAGGAGTTACCTTACCAACAAGAATATCGCCTTCCTTGACC
TCTGCACCGATACGGATAATACCCATTTCATCAAGGTCCTTGAGGGCTTCTTCA
CCAACGTTTGGCACTTCACGTGTGATTTCTTCAGGTCCA – 3'

SEQ ID n°26: partial sequence of the *rpoB* gene in *Enterococcus gallinarum* CIP 103013^T measuring 694 base pairs:

5'-
CACTCGTGAAATCCCGAATGTCGGGGAAGACGCATTGAAAGATCTAGACGAA
ATGGGTATCATCCGCATTGGTGCGGAAGTCAAAGATGGCGATCTGTTGGTTG
GTAAAGTAACGCCTAAAGGGGTAACGGAACTATCTGCAGAAGAACGCTTGCT
TCATGCAATCTTTGGTGAAAAAGCCCGCGAAGTCCGCGATACTTCTCTGCGCG
TACCTCACGGTGGTGGCGGAATCGTCCATGATGTGAAAATCTTTACCCGCGAA
GCTGGCGATGAATTGTCACCAGGTGTCAATATGCTCGTTCGCGTGTATATCGT
TCAAAAACGGAAAATCCATGAAGGGGATAAAATGGCCGGCCGTCACGGAAAT
AAAGGGGTCGTTTCTCGCATTATGCCAGAAGAAGACATGCCTTTCTTACCAGA
CGGTACACCAGTTGATATCATGTTGAACCCATTAGGGGTGCCTTCACGGATGA
ACATTGGACAAGTATTGGAATTACACTTAGGAATGGCTGCCCGCCAATTAGGA
ATCCACGTGGCTACACCAGTCTTTGATGGTGCCAGCGATGAAGATGTCTGGG
CAACAGTTGCAGAAGCCGGCATGGCTAGCGACGCCAAAACCGTTTTGTATGA
TGGCCGTACTGGAGAACCATTTGATGGTCGAATCTCCGTAGGTGTCATGTATA
TGATCAAATTGGCC– 3'

SEQ ID n°27: partial sequence of the *rpoB* gene in *Enterococcus casseliflavus* CIP 103018^T measuring 727 base pairs:

5'- TGTCATCAACCATGTGGGCCAATTTGATCATGTACATGACACCAACGGAG
ATGCGGCCATCAAATGGTTCGCCGGTACGTCCGTCGTAAAGCACTGTTTTGGC
ATCGCTGGCCATTCCTGCTTCAGCAACCGTTGCCCAAACATCTTCATCGCTGGC
TCCATCAAAGACTGGTGTTGCCACGTGAATGCCTAATTGACGCGCAGCCATTC
CTAAGTGTAACTCTAATACTTGTCCAATGTTCATCCGAGAAGGTACCCCTAATG
GGTTCAGCATGATATCGACTGGTGTGCCATCTGGTAAGAAAGGCATGTCTTCT
TCTGGCATAATGCGAGAAACGACCCCTTTGTTTCCGTGACGTCCGGCCATTTT
ATCCCCTTCATGGATTTTCCGTTTTTTGAACGATATAAACGCGAACCAGCATGTT
CACACCTGGTGACAATTCATCGCCAGCTTCGCGGGTAAAGATTTTGACATCGT
GGACGATTCCGCCGCCGCCGTGAGGCACGCGTAGAGAAGTGTCACGCACTTC
GCGGGCTTTTTCACCAAAGATTGCGTGCAACAAACGCTCTTCTGCTGAAAGTT
CCGTTACCCCTTTTGGCGTGACTTTCCCAACAAGCAGATCGCCATCTTTGACTT
CCGCACCAATGCGGATAATGCCCATTTCGTCTAGGTCTTTCAACGCGTCTTCCC
AACGTTCGGGATTTCGCGAGTGATTTCTTCAGGTCCA – 3'

SEQ ID n°28: partial sequence of the *rpoB* gene in *Enterococcus saccharolyticus* CIP 103246^T measuring 721 base pairs:

5'- TGTCATCAACCATGTGGGCAAGTTTAATCATGTACATTACCCCAACAGAG
ATACGACCATCGAATGGTTCACCCGTACGTCCGTCATAAAGAACAGTTTTCGC
ATCGCGCGCCATGCCCGCTTCGCGAACTGTTTCCCATACGTCATCATCTGATGC
ACCATCAAATACTGGTGTAGCTACATGGATGCCTAACTGACGTGCAGCCATCC
CTAAGTGTAATTCCAATACTTGTCCGATGTTCATACGAGATGGTACTCCTAGT
GGGTTCAACATGATATCAACTGGTGTGCCGTCTGGTAAGAATGGCATGTCTTC
TTCTGGCATAATGCGAGAGACAACCCCTTTGTTACCATGACGTCCCGCCATTTT
ATCTCCTTCGTGAATCTTACGTTTTTGCACGATATAAACACGAACTAACATGTT
CACACCTGGAGATAATTCGTCGCCTGCTTCACGGGTAAAGATTTTAACATCGT
GAACGATACCGCCACCGCCGTGAGGAACACGTAATGATGTATCACGTACTTCA
CGTGCTTTTTCACCGAAGATTGCGTGCAATAGACGTTCTTCTGCAGATAATTC
GGTTACCCCTTTAGGAGTGACTTTACCTACTAATAAGTCGCCATCTTGTACTTC
GGCACCGATACGGATAATACCCATTTCGTCTAAGTCTTTTAATGCGTCTTCCCC
AACGTTAGGAATTTCGCGTGTATTCTTCAG – 3'

SEQ ID n°29: partial sequence of the *rpoB* gene in *Enterococcus faecium* CIP 103014^T measuring 727 base pairs:

5'- TGTCATCAACCATGTGAGCAAGTTTGATCATGTACATCACACCGACAGAC
ACACGTCCATCAAATGGTTCACCTGTACGTCCGTCGTACAGAACAGTTTTCGC
ATCGCTGGCCATACCGGCTTCACGAACTGTTTCCCATACGTCTTCATCACTTGC
ACCATCAAATACTGGCGTTGCTACGTGGATACCTAACTGACGTGCAGCCATAC
CCAAGTGTAATTCCAATACTTGCCCGATGTTCATACGTGAAGGCACCCCTAAA
GGATTCAGCATGATATCGATTGGTGTTCCATCAGGTAGGAATGGCATATCTTC
TTCCGGCATAATACGGGATACAACCCCTTTATTTCCGTGACGACCGGCCATTTT
ATCCCCTTCATGGATTTTACGTTTTTGAACGATATAAACACGAACTAACATGTT
TACGCCTGGTGACAATTCATCTCCAGCTTCACGAGTAAAGATTTTCACATCGT
GAACGATACCGCCGCCGCCATGTGGTACACGTAATGATGTATCGCGGACTTCA
CGAGCTTTTTCGCCAAAGATCGCATGCAATAGACGTTCTTCTGCAGATAATTCT
GTTACCCCTTTTGGCGTGACTTTCCCTACAAGCAAATCGCCATCTTGGACTTCT
GCACCAATACGGATGATACCCATTTCGTCTAAATCTTTTAATGCGTCTTCCCGA
CATTAGGGATTTCGCGTGTGATTTCTTCAGGTCCA – 3'

SEQ ID n°30: partial sequence of the *rpoB* gene in *Enterococcus faecalis* CIP 103015^T measuring 724 base pairs:

5'- TGTCATCAACCATGTGGGCTAATTTAATCATATACATGACACCAACGGAA
ATACGGTTATCAAATGGTTCACCTGTACGTCCATCGTAAAGAACTGTTTTAGC
ATCGCTAGCCATACCAGCTTCACGAACAGTTTCCCAAACGTCTTCATCGGTTGC
CCCATCGAAAACAGGTGTTGCGACGTGAATACCTAATTGGCGAGCAGCCATAC
CTAAGTGTAATTCAAGTACTTGTCCGATATTCATACGAGAAGGTACCCCTAAT
GGGTTCAACATGATATCAACAGGTGTTCCGTCAGGTAAGAATGGCATATCTTC
TTCCGGCATAATACGGGAAACAACCCCTTTATTTCCGTGACGTCCCGCCATTTT
ATCTCCTTCGTGAATTTTACGTTTTTGAACGATATAGACACGAACTAACATGTT
GACACCTGGTGATAATTCATCGCCAGCTTCACGAGTAAAGATTTTCACATCAT
GAACGATACCGCCGCCACCGTGAGGTACACGGAGAGACGTATCACGAACTTC
GCGGGCTTTTTCCCCGAAGATTGCGTGTAATAAACGTTCTTCTGCAGATAATT
CTGTGACCCCTTTAGGTGTGACTTTCCCAACTAGTAAGTCGCCATCTTGAACTT
CAGCACCAATGCGGATAATCCCCATTTCGTCTAAGTCTTTCAACGCGTCTTCCC
AACGTTTGGAATTTCACGGGTATTTCTTCAGGTCA – 3'

SEQ ID n°31: partial sequence of the *rpoB* gene in *Enterococcus avium* CIP 103019^T measuring 570 base pairs:

5'- GTCCATCATAAAGAACGGTCTTAGCATCTGCTGCCATACGAGCTTCACGA
ACTGTTTCCCAAACATCGCTATCTTGCGCACCATCGAAGACTGGTGTCGCAAC
ATGGATACCTAGTTGGCGAGCCGCCATTCCCAAGTGTAATTCCAACACTTGTC
CGATGTTCATCCGAGATGGCACACCTAATGGGTTCAACATGATATCAACTGGC
GTACCGTCTGGTAAGAAAGGCATGTCTTCTTCTGGCATAATGCGAGAAACGA
CCCCTTTATTTCCGTGACGGCCGGCCATTTTATCCCCTTCATGAATCTTACGTT
TTTGCACGATGTACACGCGCACTAACATATTTACACCTGGAGATAATTCATCGC
CTGCTTCACGAGTAAAGATCTTCACATCGTGAACGATCCCGCCGCCACCATGC
GGTACACGAAGAGATGTATCACGAACTTCACGAGCCTTTTCACCAAAGATCGC
ATGCAACAAACGTTCTTCAGCTGATAATTCTGTTACCCCTTTAGGAGTGACTTT
ACCAACTAATAAATCACCATCATGAACTTCAGCACCAATAC -3'

SEQ ID n°32: partial sequence of the *rpoB* gene in *Abiotrophia defectiva* CIP 103242^T measuring 732 base pairs:

5'- GAAGTTGTCATCAACCATGTGGGCCAACTTAATCATGTACATAACCCCAA
CAGAGACTTTACGGTCAAATGGTTCACCGGTTCGACCATCATATAAGATAGTC
TTAGCGTCAGCTTCTAAGCCGGCTTCCTTAACTGTTTCCCAGACATCTTCTTCA
CTAGCACCGTCAAAGACAGGTGTTGCAATCTTGATGCCCATTTCGCGAGCAGC
CATCCCCAAGTGTAACTCTAGGACTTGCCCGATGTTCATACGGGATGGAACCC
CTAATGGGTTCAACATGATATCAACTGGGGTACCATCTGGTAAGAATGGCATA
TCTTCTTCCGGCATGATAAGGGAGACAACCCCTTTGTTACCGTGACGACCGGC
CATCTTATCCCCTTCATTGATTTTACGTTTTTGTACGATGTAGACGCGGACTAG
CTTGTTGACACCTGGTGCCAATTCGTCGCCAGCTTCGCGGGTAAAGATTTTAA
CGTCGTGGACAATCCCGCCCCCGCCGTGTGGCACACGCAAGGAAGTATCACG
TACTTCACGCGCCTTCTCACCGAAGATAGCATGGAGCAAGCGTTCTTCCGCAG
ACAACTCGGTCACACCTTTTGGTGTTACCTTACCAACTAAGATATCGCCGTCTT
TTACTTCCGCCCCGATACAGATAATCCCGTCTTGGTCTAAGTACTTGAGGGCA
TCTTCGGACACGTTTGGAATTTCGCGTGTAATTTCTTCAGGTCA – 3'

SEQ ID n°33: partial sequence of the *rpoB* gene in *Gemella morbilorum* CIP 81.10^T measuring 727 base pairs:

5'- TGTCATCAACCATGTGTGCAAGTTTATCATGTACATTACCCCTACAGATAC
ACGGCTATCAAATGGCTCACCTGTACGTCCGTCATAAAGAACTGTCTTAGCAT
CTTTAGCCATTCCAGCTTCCGCAACTGTAGACCAAACATCTTCATCAGTAGCAC
CATCGAATACTGGTGTAGCTACGTGGATTCCAAGTTGTTTAGCAGCCATACCT
AAGTGTAGCTCTAATACTTGTCCAATGTTCATACGAGATGGAACCCCAAGTGG
GTTTAACATTACGTCAACTGGTGTACCATCTGGTAGGTAAGGCATATCTTCTT
CTGGTAAGATATTTGAGATAACCCCTTTGTTACCGTGACGACCGGCCATTTTA
TCTCCTACACGAATTTTACGTTTTTGGACGATAAATACACGAACAAGTTCATTT
ACACCGTTAGGTAATTCAGCACCATCTTCACGTTTAAAGATTTTAACATCAGCA
ACTACTCCATCAGCACCGTGAGGTACACGTAATGAAGTATCACGTACTTCTTTA
GATTTAGCTCCAAAGATAGCATATAATAATTTTTCTTCTGGAGTTTGTTCAGTT
AATCCTTTCGGTGTAACTTTACCTACTAAAATATCTCCATCTTTAACTTCAGCC
CCAATACGAATGATTCCTCGTGCATCTAAGTTTCTAAGTGCATTTTCACCCTAC
GTTTGGAATCTCACGAGTAATTTCTTCAGGTCA – 3'

SEQ ID n°34: partial sequence of the *rpoB* gene in *Gemella haemolysans* CIP 101126^T measuring 726 base pairs:

5'- TGTCATCAACCATGTGTGCAAGTTTAATCATGTACATTACCCCTACAGATA
CACGGCTATCAAATGGCTCACCTGTACGTCCGTCATAAAGAACTGTCTTAGCA
TCTTTAGCCATTCCAGCTTCCGCAACTGTAGACCAAACATCTTCATCAGTAGCA
CCATCGAATACTGGTGTAGCTACGTGGATTCCAAGTTGTTTAGCAGCCATACC
TAAGTGTAGCTCTAATACTTGTCCAATGTTCATACGAGATGGAACCCCAAGTG
GGTTTAACATTACGTCAACTGGTGTACCATCTGGTAGGTAAGGCATATCTTCT
TCTGGTAAGATATTTGAGATAACCCCTTTGTTACCGTGACGACCGGCCATTTT
ATCTCCTACACGAATTTTACGTTTTTGGACGATAAATACACGAACAAGTTCATT
TACACCGTTAGGTAATTCAGCACCATCTTCACGTTTAAAGATTTTAACATCAGC
AACTACTCCATCAGCACCGTGAGGTACACGTAATGAAGTATCACGTACTTCTTT
AGATTTAGCTCCAAAGATAGCATATAATAATTTTTCTTCTGGAGTTTGTTCAGT
TAATCCTTTCGGTGTAACTTTACCTACTAAAATATCTCCATCTTTAACTTCAGC
CCCAATACGAATGATTCCTCGTGCATCTAAGTTTCTAAGTGCATTTTCACCTAC
GTTTGGAATCTCACGAGTATTCTTCAGGTCCA – 3'

SEQ ID n°35: partial sequence of the *rpoB* gene in *Granulicatella adjacens* CIP 103243^T measuring 719 base pairs:

5'- CATCAACCATGTGAGCAAGTTTGATCATGTACATAACCCCTACTGACACA
CGGTTATCGAATGGTTCCCCTGTACGTCCATCATATAGAATTGTTTTCGCATCA
CGAGCCATACCCGCTTCTGCAACAGTTCCCCATACGTCTTCATCTTGCGCACCA
TCGAATACTGGTGTTGCGATGTAAATACCTAATTCACGAGCAGCCATCCCTAA
GTGTAACTCTAACACTTGTCCGATGTTCATACGTGAAGGTACCCCTAATGGGT
TTAACATGATGTCAACTGGTGTTCCATCTGGTAAGAATGGCATATCTTCTTCC
GGCATAATACGGGAAACAACCCCTTTATTACCGTGACGTCCGGCCATCTTATC
CCCTTCATTGATTTTACGTTTTTGTACAATATATACACGAACTAATTTGTTTACG
CCAGGTGCTAATTCATCACCTGCTGCACGTGTGAATACACGTACATCACGGAC
AATACCGCCACCGCCGTGAGGTACACGTAGAGATGTGTCACGAACTTCACGA
GCTTTTTCACCGAAGATTGCGTGTAATAAACGTTCCTCTGGTGATTGTTCTGTT
AACCCTTTAGGAGTTACTTTACCAACTAAGATGTCACCATCTTTAACTTCGGCA
CCGATACGAATAATTCCGTCTGCGTCTAGGTTCTTCAATGCGTCTTCCCAACGT
TTGGAATCTCACGAGTAATTCTTCAGG – 3'

In the above sequences, the M nucleotide designates A or C, the R nucleotide designates A or G, the W nucleotide designates A or T, the Y nucleotide designates C or T and the N nucleotide designates A, T, C or G.

In the above sequences, the CIP references relate to deposits with the national collection of microorganism cultures: *Collection Nationale de Culture des Microorganismes* (CNCM) at Institut Pasteur in Paris (France).

Example 3: Blind identification of a collection of 20 bacterial strains comprising 10 strains of bacteria belonging to genus *Streptococcus* and related genera.

A collection of twenty strains belonging to the following bacterial species: *Streptococcus pyogenes, Streptococcus sanguis, Granulicatella adjacens, Abiotrophia defectiva, Enterococcus avium, Enterococcus faecalis, Gemella*

haemolysans, Gemella morbilorum, Streptococcus equi, Streptococcus anginosus, Staphylococcus aureus, Pseudomonas oleovorans, Mycobacterium avium, Bacillus cereus, Acinetobacter anitratus, Corynebacterium amycolatum, Klebsiella terrigena, Pasteurella, Lactobacillus rhamnosus, Staphylococcus was coded so as to conduct blind molecular identification of strains (the experimenter not having any *a priori* knowledge of strain identity) using the method described in the present patent application. Extraction of the nucleic acids and amplification of the *rpoB* gene fragment were performed as described in example 2 incorporating primers consisting of mixtures of 4 oligonucleotides which have sequences consisting of sequences SEQ ID n°6 (as 5′ primer) and SEQ ID n°7 (as 3′ primer) where N represents inosine, in a PCR amplification (Fig.1). The sequencing of these 10 amplificates was conducted by incorporating into the sequencing reaction the primers SEQ ID n° 6 and SEQ ID n° 7 as described in example 2, and comparison of the sequences obtained with sequences SEQ ID n° 1 to 5 and 8 to 35 enabled the 10 ten amplified strains to be identified as being *Streptococcus pyogenes, Streptococcus sanguis, Granulicatella adjacens, Abiotrophia defectiva, Enterococcus avium, Enterococcus faecalis, Gemella haemolysans, Gemella morbilorum, Streptococcus equi, Streptococcus anginosus*. The decoding of these 10 strains showed 100% agreement between molecular identification using the method that is the subject of the invention and the identification previously established by standard phenotype methods. This result illustrates the specificity of the set of primers SEQ ID n°6/SEQ ID n°7.

The other bacteria chosen because they are frequently isolated in human or animal clinical specimens and also possibly contain bacteria of genus *Streptococcus* were not amplified, thereby exhibiting the specificity of the primers

used for the *Streptococcus* genus and said 4 related genera under the conditions of use of the invention for detecting bacteria of genus *Streptococcus* and said 4 related genera in comparison with bacteria of another genus.

Figure 1 shows the PCR amplification products obtained from ten coded bacterial strains, comprising 7 strains belonging to genus *Streptococcus* and said 4 related genera (columns 2,3,4, 7-11) and 3 bacterial strains of bacterial genera other than *Streptococcus* and said 4 related genera (columns 5, 6 and 12). Columns 1 and 13 show the molecular weight marker. The amplification products are obtained after incorporating primers SEQ ID n° 6 and SEQ ID n° 7 described above, and are visualized by staining with ethidium bromide after electrophoresis on agarose gel.